

## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:44:23 ; Search time 136 Seconds  
(without alignments)

951.880 Million cell updates/sec

Title: US-10-031-918A-1

Perfect score: 1997

Sequence: 1 MVSVSIRKAQRAEGPATIL.....LFGFGPLTETVLSVAI 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	99.8	389	16	US-10-450-230-1
2	1994	99.8	389	16	US-10-450-230-10
3	1969	98.6	389	16	US-10-659-755-7
4	1933	96.8	389	17	US-10-491-823-316
5	1911	95.7	389	16	US-10-659-755-9
6	1903	95.3	389	17	US-10-491-823-314
7	1840	92.1	389	17	US-10-491-823-320
8	1839	92.1	405	15	US-10-425-114-52518
9	1820	91.1	386	15	US-10-424-599-148762
10	1809.5	90.6	388	15	US-10-424-599-171157
11	1808.5	90.6	383	15	US-10-424-599-171068
12	1804.5	90.4	388	15	US-10-424-599-171091
13	1803.5	90.3	388	15	US-10-424-599-171202

14	1798.5	90.1	388	15	US-10-424-599-171113	Sequence 171113,
15	1795.5	89.9	388	15	US-10-424-599-171124	Sequence 171124,
16	1761	88.2	389	16	US-10-641-495-5	Sequence 5, Appli
17	1734	86.8	389	14	US-10-136-444-2	Sequence 2, Appli
18	1733	86.8	391	15	US-10-424-599-169160	Sequence 169160,
19	1719	86.1	387	15	US-10-425-114-50078	Sequence 50078, A
20	1715.5	85.9	384	15	US-10-424-599-148763	Sequence 148763,
21	1705.5	85.4	396	16	US-10-450-230-9	Sequence 9, Appli
22	1696	84.9	392	15	US-10-224-493-2	Sequence 2, Appli
23	1692	84.7	392	15	US-10-224-493-6	Sequence 6, Appli
24	1690	84.6	392	15	US-10-224-493-4	Sequence 4, Appli
25	1686.5	84.5	397	15	US-10-425-114-40814	Sequence 40814, A
26	1676.5	84.0	401	15	US-10-425-114-37280	Sequence 37280, A
27	1672.5	83.8	390	15	US-10-424-599-171190	Sequence 171190,
28	1671	83.7	391	17	US-10-491-823-324	Sequence 324, App
29	1660.5	83.1	395	16	US-10-715-872-168	Sequence 168, App
30	1634	81.8	398	16	US-10-437-963-165957	Sequence 165957,
31	1630	81.6	394	14	US-10-289-757-119	Sequence 119, App
32	1630	81.6	398	14	US-10-289-757-186	Sequence 186, App
33	1630	81.6	398	15	US-10-431-273-76	Sequence 76, Appl
34	1627	81.5	398	15	US-10-431-273-77	Sequence 77, Appl
35	1550	77.6	309	17	US-10-491-823-24	Sequence 24, Appl
36	1544	77.3	360	9	US-09-829-378-3	Sequence 3, Appli
37	1543	77.3	382	16	US-10-437-963-165955	Sequence 165955,
38	1524	76.3	338	15	US-10-425-114-43099	Sequence 43099, A
39	1508.5	75.5	394	9	US-09-837-654-1	Sequence 1, Appli
40	1508.5	75.5	394	14	US-10-390-658-1	Sequence 1, Appli
41	1492	74.7	398	16	US-10-767-701-45589	Sequence 45589, A
42	1439.5	72.1	403	16	US-10-437-963-148991	Sequence 148991,
43	1409.5	70.6	393	16	US-10-450-230-8	Sequence 8, Appli
44	1293	64.7	360	9	US-09-829-378-4	Sequence 4, Appli
45	1243	62.2	404	16	US-10-767-701-46457	Sequence 46457, A

#### ALIGNMENTS

#### RESULT 1

US-10-450-230-1  
; Sequence 1, Application US/10450230  
; Publication No. US20040106175A1  
; GENERAL INFORMATION:  
; APPLICANT: NOEL, JOSEPH P.  
; APPLICANT: AUSTIN, MICHAEL B.  
; APPLICANT: BOWMAN, MARIANNE E.  
; TITLE OF INVENTION: METHODS OF PRODUCING POLYKETIDE SYNTHASE MUTANTS AND  
; TITLE OF INVENTION: COMPOSITIONS AND USES THEREOF  
; FILE REFERENCE: SALK2370-3  
; CURRENT APPLICATION NUMBER: US/10/450,230  
; CURRENT FILING DATE: 2003-11-26  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR FILING DATE: 2000-12-15  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Medicago sativa  
US-10-450-230-1

Query Match 99.8%; Score 1994; DB 16; Length 389;  
Best Local Similarity 99.7%; Pred. No. 7.2e-192;  
Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVSIRKAQRAEGPATILAITGTANPANCVCQSTYDFYFKITNSEHKTKEKQFQMC 60

Db 1 MVSVSIRKAQRAEGPATILAITGTANPANCVCQSTYDFYFKITNSEHKTKEKQFQMC 60

QY 61 DKSMIKRRYYMLTEELKENPNVCYMAPSLDARQDMVVVPRLGKAAVKAKEWGQP 120

Db 61 DKSMIKRRYYMLTEELKENPNVCYMAPSLDARQDMVVVPRLGKAAVKAKEWGQP 120

Qy	121	KS	K	S	K	I	T	H	L	I	V	C	T	T	S	G	V	D	M	P	G	A	D	Y	Q	L	T	K	L	G	L	R	P	Y	V	K	R	Y	M	Y	Q	O	G	F	A	G	T	V	L	R	L	A	O	A	E	N	180			
Db	121	KS	K	S	K	I	T	H	L	I	V	C	T	T	S	G	V	D	M	P	G	A	D	Y	Q	L	T	K	L	G	L	R	P	Y	V	K	R	Y	M	Y	Q	O	G	F	A	G	T	V	L	R	L	A	O	A	E	N	180			
Qy	181	NG	K	A	R	V	L	V	V	C	S	E	V	T	A	T	F	R	G	P	S	D	T	H	L	S	I	S	G	A	L	F	G	D	A	A	L	I	V	G	S	P	P	V	E	I	K	P	I	F	E	M	240							
Db	181	NG	K	A	R	V	L	V	V	C	S	E	V	T	A	T	F	R	G	P	S	D	T	H	L	S	I	S	G	A	L	F	G	D	A	A	L	I	V	G	S	P	P	V	E	I	K	P	I	F	E	M	240							
Qy	241	W	T	A	Q	T	I	A	P	S	E	G	A	I	D	G	H	L	R	E	A	G	L	T	P	H	L	L	K	D	V	P	G	I	V	S	K	N	I	T	K	A	L	V	E	A	P	E	L	G	I	S	D	Y	N	S	I	F	W	300
Db	241	W	T	A	Q	T	I	A	P	S	E	G	A	I	D	G	H	L	R	E	A	G	L	T	P	H	L	L	K	D	V	P	G	I	V	S	K	N	I	T	K	A	L	V	E	A	P	E	L	G	I	S	D	Y	N	S	I	F	W	300
Qy	301	I	A	H	P	G	P	A	L	D	O	V	E	K	L	A	K	P	E	K	N	A	T	R	E	V	L	S	E	G	N	M	S	S	A	C	V	L	F	I	L	D	E	M	R	K	S	T	O	N	G	L	K	T	360					
Db	301	I	A	H	P	G	P	A	L	D	O	V	E	K	L	A	K	P	E	K	N	A	T	R	E	V	L	S	E	G	N	M	S	S	A	C	V	L	F	I	L	D	E	M	R	K	S	T	O	N	G	L	K	T	360					
Qy	361	T	G	S	E	L	W	G	V	L	F	G	F	G	L	T	I	E	T	V	L	R	S	V	A	I	389																																	
Db	361	T	G	S	E	L	W	G	V	L	F	G	F	G	L	T	I	E	T	V	L	R	S	V	A	I	389																																	

## RESULT 2

```

US-10-450-230-10
; Sequence 10, Application US/10450230
; Publication No. US20040106175a1
; GENERAL INFORMATION:
; APPLICANT: NOEL, JOSEPH P.
; APPLICANT: AUSTIN, MICHAEL B.
; APPLICANT: BOWMAN, MARRIANNE E.
; TITLE OF INVENTION: METHODS OF PRODUCING POLYKETIDE SYNTHASE MUTANTS AND
; FILE REFERENCE: COMPOSITIONS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/450,230
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: PCT/US01/48523
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/255,811
; PRIOR FILING DATE: 2000-12-15
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Medicago sativa
US-10-450-230-10

```

```

Db      301  IAHGGPAILDQVBKQALKPEKNATREVLSEYGNMSSACVLFILDEMRRKKTQNGLKT 360
Qy      361  TGELEWGLVFGFGPGLTIETVWLSVAI 389
          |||||
Db      361  TGELEWGLVFGFGPGLTIETVWLSVAI 389
          |||||

RESULT 3
US-10-659-755-7
; Sequence 7, Application US/10659755
; Publication No. US20040128711A1
; GENERAL INFORMATION:
; APPLICANT: DIXON, RICHARD A.
; APPLICANT: LIU, CHANG-JUN
; APPLICANT: DEAVOURS, BETTINA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PRODUCTION OF FLAVONOID
; TITLE OF INVENTION: AND ISOFLAVONOID NUTRACEUTICALS
; FILE REFERENCE: NBLE:007US
; CURRENT APPLICATION NUMBER: US/10/659,755
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/409,447
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Medicago sativa
US-10-659-755-7

```

RESULT 4

```

RESULT 4
US-10-491-823-316
Sequence 316, Application US/10491823
; Publication No. US2005069884A1
; GENERAL INFORMATION:
; APPLICANT: Agriculture Victoria Services Pty Ltd
; APPLICANT: Agresearch Limited
; APPLICANT: SPANGENBERG, German
; APPLICANT: SAWBRIDGE, Timothy Ivor
; APPLICANT: ONG, Eng Kok

```

```

; APPLICANT:  EMMERLING, Michael
; TITLE OF INVENTION: Manipulation of flavonoid biosynthesis in plants
; FILE REFERENCE: 4491-4003
; CURRENT APPLICATION NUMBER: US/10/491.823
; CURRENT FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: PR8113
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: PCT/AU2002/001345
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 477
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 389
; TYPE: PR1
; ORGANISM: Trifolium repens
; US-10-491-823-316

Query Match          96.8%; Score 1933; DB 17; Length 389;
Best Local Similarity 96.7%; Pred. No. 1e-185;
Matches 376; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MVSVEIRKQRAESGPATILAIGTANPANCVEQSTYPDFYFKITNSEHKTELKEKFORMC 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MVSVAEIRKQRAESGPATILAIGTANPPNRVEQSTYPDFYFKITNSEHKTELKEKFORMC 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 DKSMIKERYMYLTHEILKENPNVCEYMAPSLDARQDMVVVPPRLGKEAAVKAKEWQOP 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DKSMIKERYMYLTHEILKENPSLCEYMAPSLDARQDMVVVPPRLGKEAAVKAKEWQOP 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 121 KSKITHLIVICTSGVDMPGADYQUTKLLGLRPYVKRYMYQGXFAGGTVLRLAKDLAEN 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 KSKITHLIFCTTSGVDMPGADYQUTKLLGLRPYVKRYMYQGXFAGGTVLRLAKDLAEN 180
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 181 NKGARVLVVCSEVTAVTFRGSPTHLSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 NKGARVLVVCSEVTAVTFRGSPTHLSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAPEPIGIDSYNSIFW 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNINKALVEAFQPLGIDSYNSIFW 300
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLPILDEMRRKKSQNGLKT 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLPILDEMRRKKSQNGLKT 360
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 361 TGEGLWGVLFPGFGPLTIETVLRVSAI 389
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 TGEGLDWGVLFPGFGPLTIETVLRVSAI 389
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-10-659-755-9
; Sequence 9, Application US/10659755
; Publication No. US20040128711A1
; GENERAL INFORMATION:
; APPLICANT: DIXON, RICHARD A.
; APPLICANT: LIU, CHANG-JUN
; APPLICANT: DEAVOURS, BETTINA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR PRODUCTION OF FLAVONOID
; TITLE OF INVENTION: AND ISOFLAVONOID NUTRACEUTICALS
; FILE REFERENCE: NBLE:0070US
; CURRENT APPLICATION NUMBER: US/10/659,755
; CURRENT FILING DATE: 2003-09-10
; PRIOR APPLICATION NUMBER: 60/409,447
; PRIOR FILING DATE: 2002-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 389
; TYPE: PR1
; ORGANISM: Medicago sativa
; US-10-659-755-9

```

Db 121 KSKITHLIFCTTSGVDMFGADYQLTKLLGLRPVVKRYMYQQCGFAGGTVLRLAKDLAEN 180  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 240  
Db 181 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
Qy 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389

## RESULT 7

US-10-491-823-320  
; Sequence 320, Application US/10491823  
; Publication No. US20050069884A1  
; GENERAL INFORMATION:  
; APPLICANT: Agriculture Victoria Services Pty Ltd  
; APPLICANT: AgResearch Limited  
; APPLICANT: SPANGENBERG, German  
; APPLICANT: SAWBRIDGE, Timothy Ivor  
; APPLICANT: ONG, Eng Kok  
; APPLICANT: EMMERLING, Michael  
; TITLE OF INVENTION: Manipulation of flavonoid biosynthesis in plants  
; FILE REFERENCE: 4491-4003  
; CURRENT APPLICATION NUMBER: US/10/491,823  
; CURRENT FILING DATE: 2004-04-05  
; PRIOR FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: PCT/AU2002/001345  
; PRIOR FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 477  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 320  
; TYPE: PRT  
; ORGANISM: Trifolium repens  
US-10-491-823-320

Query Match 92.1%; Score 1840; DB 17; Length 389;  
Best Local Similarity 89.2%; Pred. No. 2.4e-176;  
Matches 346; Conservative 28; Mismatches 14; Indels 0; Gaps 0;  
Qy 1 MVSVEIRKAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
Db 1 MVSVEIRKAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
Qy 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRLCKEAAVKAKEWGQP 120  
Db 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRLCKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMFGADYQLTKLLGLRPVVKRYMYQQCGFAGGTVLRLAKDLAEN 180  
Db 121 KSKITHLIVCTTSGVDMFGADYQLTKLLGLRPVVKRYMYQQCGFAGGTVLRLAKDLAEN 180  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 240  
Db 181 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
Qy 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

Db 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
RESULT 8  
US-10-425-114-52518  
; Sequence 52518, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 52518  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3049-050-G10\_FLI.pap  
US-10-425-114-52518

Query Match 92.1%; Score 1839; DB 15; Length 405;  
Best Local Similarity 90.2%; Pred. No. 3.2e-176;  
Matches 351; Conservative 19; Mismatches 19; Indels 0; Gaps 0;  
Qy 1 MVSVEIRKAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
Db 17 MVSVAEIRQAQRAEGPATILAITANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 76  
Qy 61 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRLCKEAAVKAKEWGQP 120  
Db 77 DKSIMKRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRLCKEAAVKAKEWGQP 136  
Qy 121 KSKITHLIVCTTSGVDMFGADYQLTKLLGLRPVVKRYMYQQCGFAGGTVLRLAKDLAEN 180  
Db 137 KSKITHLIVCTTSGVDMFGADYQLTKLLGLRPVVKRYMYQQCGFAGGTVLRLAKDLAEN 196  
Qy 181 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 240  
Db 197 NKGARVLVVCSEVTANTFRGSDTHLSDLVGOALFGDGAALIVGSDPVEIEKPIFEMV 256  
Qy 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
Db 257 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 316  
Qy 301 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
Db 317 IAHGGPAILDQVEQKALPEKONATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 376  
Qy 361 TGEGLWGLVFGFGPGLTIETVLRSAI 389  
Db 377 TGEGLWGLVFGFGPGLTIETVLRSAI 405

## RESULT 9

US-10-424-599-148762  
; Sequence 148762, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 148762  
LENGTH: 386  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_105354C.1.pap  
US-10-424-599-148762

Query Match 91.1%; Score 1820; DB 15; Length 386;  
Best Local Similarity 90.8%; Pred. No. 2.5e-174; Indels 0; Gaps 0;  
Matches 344; Conservative 19; Mismatches 16; Indels 0; Gaps 0;  
QY 1 MVSVEIRKAQRAEGPATILAICTANPANCVEQSTYDPDYFKITNSEHKTTELKEKFRMC 60  
DB 1 MVSVAEIRKAQRAEGPATILAICTANPPNCVQSTYDPDYFRITNSEHMTTELKEKFRMC 60  
QY 61 DKSMIKRRYMYLTEELKENPNVCYMAPSLDARQDMVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLNEELKENPNVCAYMAPSLDARQDMVVEVPKLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGXFAGGTVLRLLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGCFAGGTVLRLLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEITAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGIVSKNIDKALFEAFNPLINISDYNSIFW 300  
QY 301 IAHGPGPALDQVEOKLALPKPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT 360  
DB 301 IAHGPGPALDQVEOKLALPKPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT 360  
QY 361 TGEGLWGVLFGRPGGLTI 379  
DB 361 TGEGLWGVLFGRPGGLTL 379

RESULT 10  
US-10-424-599-171157  
Sequence 171157, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 171157  
LENGTH: 388  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12556C.1.pap  
US-10-424-599-171157

Query Match 90.6%; Score 1809.5; DB 15; Length 388;  
Best Local Similarity 87.9%; Pred. No. 2.8e-173;

Matches 342; Conservative 26; Mismatches 20; Indels 1; Gaps 1;  
QY 1 MVSVEIRKAQRAEGPATILAICTANPANCVEQSTYDPDYFKITNSEHKTTELKEKFRMC 60  
DB 1 MVSVEIRKAQRAEGPATVMAIGTATPNCVQSTYDPDYFRITNSEHMTTELKEKFRMC 60  
QY 61 DKSMIKRRYMYLTEELKENPNVCYMAPSLDARQDMVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLNEELKENPNVCAYMAPSLDARQDMVVEVPKLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGXFAGGTVLRLLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGCFAGGTVLRLLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEITAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVP-VEKPLFQLV 239  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREVGLTFHLLKDVPGISKNIEKALVEAFQPLIGSDYNSIFW 299  
QY 301 IAHGPGPALDQVEOKLALPKPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT 360  
DB 301 IAHGPGPALDQVEAKLGKPKMEATRHVLSEYGNMSSACVLFILDQMRKKSIEGLGT 359  
QY 361 TGEGLWGVLFGRPGGLTIETVLRSAI 389  
DB 361 TGEGLDMGVLFGRPGGLTIVETVLRSAITV 388

RESULT 11  
US-10-424-599-171068  
Sequence 171068, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 171068  
LENGTH: 583  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12548C.1.pap  
US-10-424-599-171068

Query Match 90.6%; Score 1808.5; DB 15; Length 583;  
Best Local Similarity 87.9%; Pred. No. 6.8e-173;  
Matches 342; Conservative 26; Mismatches 20; Indels 1; Gaps 1;  
QY 1 MVSVEIRKAQRAEGPATILAICTANPANCVEQSTYDPDYFKITNSEHKTTELKEKFRMC 60  
DB 196 MVSVEIRKAQRAEGPATVMAIGTATPNCVQSTYDPDYFRITNSEHMTTELKEKFRMC 255  
QY 61 DKSMIKRRYMYLTEELKENPNVCYMAPSLDARQDMVVEVPRLGKEAAVKAKEWGQP 120  
DB 256 DKSMIKRRYMYLNEELKENPNVCAYMAPSLDARQDMVVEVPKLGKEAAVKAKEWGQP 315  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGXFAGGTVLRLLAKDLAEN 180  
DB 316 KSKITHLIFCTTSGVDMPGADYQTLKGLRPPVVKRYMYQQGCFAGGTVLRLLAKDLAEN 375  
QY 181 NKGARVLVVCSEVTAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
DB 376 NKGARVLVVCSEITAVTFRGPSTDLHSLVGQALFGDGAALIVGSDPVP-VEKPLFQLV 434





OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12552C.1.pep  
US-10-424-599-171113

Query Match 90.1%; Score 1798.5; DB 15; Length 388;  
Best Local Similarity 87.4%; Pred. No. 3.7e-172; Indels 1; Gaps 1;  
Matches 340; Conservative 26; Mismatches 22; Indels 1; Gaps 1;  
  
QY 1 MVSVEIRKAQRAEGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATVMAIGTATPPNCVDQSTYDPDYFRITNSEHMTTELKEKFKRMC 60  
  
QY 61 DKSMIKRYYMYLTEEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
DB 61 DKSMIKKLYMYLNEEILKENPSYCAYNAPSLDARDQMVVVEPRLGKEAATKAKEWGQP 120  
  
QY 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPCGADYQTLKLLGLRPPSVKRYMYQQGCFAGGTVLRLAKDLAEN 180  
  
QY 181 NKGARVLVVCSEVTAVTFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEITAVTFRGPTDTHLDSLVGQALFGDGAALIVGSDPLP-VEKPLFQLI 239  
  
QY 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIUSKNITKALVEAFEPGLGSDYNSIFW 300  
DB 240 WTAQTILPDSEGAIDGHLREVGLTFHLLKDVPGLSKNIEKALVEAFQPLGSDYNSIFW 299  
  
QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT 360  
DB 300 IAHPGGPAILDQVEAKLGLKPEKMEATRHVLSEYGNMSSACVLFILDMRKKSIENGLGT 359  
  
QY 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389  
DB 360 TGEGLDWGVLFPGFGLTIVETVVLRSVTL 388

RESULT 15

US-10-424-599-171124  
; Sequence 171124, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 171124  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12553C.1.pep  
US-10-424-599-171124

Query Match 89.9%; Score 1795.5; DB 15; Length 388;  
Best Local Similarity 87.7%; Pred. No. 7.3e-172; Indels 1; Gaps 1;  
Matches 341; Conservative 26; Mismatches 21; Indels 1; Gaps 1;  
  
QY 1 MVSVEIRKAQRAEGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATVMAIGTATPPNCVDQSTYDPDYFRITNSEHMTTELKEKFKRIC 60  
  
QY 61 DKSMIKRYYMYLTEEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAIKEWGQP 120  
DB 61 DKSMIKKRYMYLNEEILKENPSYCAYNAPSLDARDQMVVVEPRLGKEAATKAKEWGQP 120  
  
QY 121 KSKITHLIVCTTSGVDMPCGADYQTLKLLGLRPPVVKRYMYQQGXFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPCGADYQTLKLLGLRPPSVKRYMYQQGCFAGGTVLRLAKDLAEN 180

DB 121 KSKITHLIFCTTSGVDMPCGADYQTLKLLGLRPPSVKRYMYQQGCFAGGTVLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPIEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEITAVTFRGPTDTHLDSLVGQALFGDGAALIVGSDPLP-VEKPLFQLV 239  
QY 241 WTAQTAPDSEGAIDGHLREAGLTTHLLKDVPGIUSKNITKALVEAFEPGLGSDYNSIFW 300  
DB 240 WTAQTILPDSEGAIDGHLREVGLTFHLLKDVPGLSKNIEKALVEAFQPLGSDYNSIFW 299  
QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRRKKSSTONGLKT 360  
DB 300 IAHPGGPAILDQVEAKLGLKPEKMEATRHVLSEYGNMSSACVLFILDMRKKSIENGLGT 359  
QY 361 TGEGLWGVLFPGFGLTIETVVLRSVAI 389  
DB 360 TGEGLDWGVLFPGFGLTIVETVVLRSVTL 388

Search completed: April 22, 2005, 21:57:07  
Job time : 138 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:27:00 ; Search time 172 Seconds  
(without alignments)  
1158.133 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVEIRKAQRAEGPATIL.....LFGFGGLTIETVLRVAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*

1: uniprot\_prot.\*

2: uniprot\_tmbl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1994	99.8	389	1 CHS2_MEDSA	P30074 medicago sa
2	1969	98.6	389	1 CHS4_MEDSA	P30075 medicago sa
3	1966	98.4	389	2 Q9FEW4	Q9FEW4 medicago tr
4	1941	97.2	389	1 CHS8_MEDSA	P30076 medicago sa
5	1938	97.0	389	1 CHS5_MEDSA	P51078 medicago sa
6	1937	97.0	389	1 CHS5_PEA	Q23884 pisum sativ
7	1932	96.7	389	1 CHS2_PEA	Q01287 pisum sativ
8	1930	96.6	389	1 CHS3_PEA	O23883 pisum sativ
9	1927	96.5	389	1 CHS2_TRISU	P51084 trifolium s
10	1926	96.4	389	1 CHS1_PEA	Q01286 pisum sativ
11	1926	96.4	389	1 CHS9_MEDSA	P30077 medicago sa
12	1925	96.4	389	1 CHS4_PEA	O23882 pisum sativ
13	1920	96.1	389	1 CHS6_TRISU	P51088 trifolium s
14	1915	95.9	389	1 CHS6_PEA	Q01288 pisum sativ
15	1914	95.8	389	1 CHS5_TRISU	P51087 trifolium s
16	1912	95.7	389	1 CHS3_TRISU	P51085 trifolium s
17	1911	95.7	389	1 CHS1_MEDSA	P51077 medicago sa
18	1908	95.5	389	1 CHS3_TRISU	P51083 trifolium s
19	1908	95.5	389	2 Q9ZRV7	Q9ZRV7 cicer ariet
20	1903	95.3	389	1 CHS1_CICAR	Q9SM14 cicer ariet
21	1881.5	94.2	390	1 CHS7_ONOVI	O22586 onobrychis
22	1880	94.1	389	1 CHS1_MEDSA	P30073 medicago sa
23	1859	93.6	389	2 Q84V87	Q84V87 glycine max
24	1857	93.5	389	1 CHS7_PUELO	P23569 pueraria lo
25	1854	92.8	389	2 Q8S4Y7	Q8S4Y7 senna alata
26	1844	92.3	389	1 CHS7_PHAVU	P49440 phaseolus v
27	1839	92.1	389	1 CHS7_SOYBN	P30081 glycine max
28	1838	92.0	389	2 Q84XJ8	Q84XJ8 arachis hyp
29	1825	91.4	389	2 Q8S4Y6	Q8S4Y6 senna alata
30	1818	91.0	389	1 CHS4_PEA	P51081 pisum sativ
31	1817	91.0	389	2 Q8S4Y8	Q8S4Y8 senna alata

32	1814	90.8	389	1 CHS8_PEA	P51082 pisum sativ
33	1811.5	90.7	388	1 CHS1_SOYBN	P24826 glycine max
34	1809.5	90.6	388	2 Q6XON0	Q6XON0 glycine max
35	1807.5	90.5	388	2 Q6XOM9	Q6XOM9 glycine max
36	1805.5	90.4	388	1 CHS5_SOYBN	P48406 glycine max
37	1800.5	90.2	388	1 CHS3_SOYBN	P19468 glycine max
38	1800.5	90.2	388	1 CHS1_VIGUN	P51089 vigna ungui
39	1797	90.0	389	2 Q8L6D2	Q8L6D2 rhododendro
40	1795.5	89.9	388	1 CHS2_SOYBN	P17957 glycine max
41	1794.5	89.9	388	1 CHS6_SOYBN	P30080 glycine max
42	1794	89.8	389	1 CHS1_CASGL	Q9ZVR8 casuarina g
43	1789	89.6	391	2 Q8L5N6	Q8L5N6 malus domes
44	1782	89.2	389	2 Q8W3P6	Q8W3P6 vitis vinif
45	1781	89.2	390	2 Q9FUB7	Q9FUB7 hypericum a

## ALIGNMENTS

### RESULT 1

ID	CHS2_MEDSA	STANDARD	PRT	389 AA
AC	P30074			
DT	01-APR-1993	(Rel. 25, Created)		
DT	01-APR-1993	(Rel. 25, Last sequence update)		
DT	25-OCT-2004	(Rel. 45, Last annotation update)		
DE	Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).			
GN	Name=CHS2;			
OS	Medicago sativa (Alfalfa).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.			
OX	NCBI_TaxID=3879;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93283629; PubMed=8507827;			
RA	Jungthans H., Dalkin K., Dixon R.A.;			
RT	"Stress responses in alfalfa (Medicago sativa L.). 15.			
RT	Characterization and expression patterns of members of a subset of the			
RT	chalcone synthase multigene family.";			
RL	Plant Mol. Biol. 22:239-253(1993).			
RN	[2]			
RP	X-RAY CRYSTALLOGRAPHY (1.69 ANGSTROMS).			
RX	MEDLINE=20120577; PubMed=10653632; DOI=10.1021/bi991489f;			
RA	Jez J.M., Ferrer J.L., Bowman M.E., Dixon R.A., Noel J.P.;			
RT	"Dissection of malonyl-coenzyme A decarboxylation from polyketide			
RT	formation in the reaction mechanism of a plant polyketide synthase.";			
RL	Biochemistry 39:890-902(2000).			
CC	-!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-			
CC	tetrahydroxychalcone (also termed naringenin-chalcone or Chalcone)			
CC	which can under specific conditions spontaneously isomerize into			
CC	naringenin.			
CC	-!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA +			
CC	naringenin-chalcone + 3 CO(2).			
CC	-!- PATHWAY: Part of the biosynthetic pathway for all classes of			
CC	flavonoids, a large class of secondary plant metabolites, many of			
CC	which are brightly colored.			
CC	-!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; L02902; AAA02824.1; -			
DR	PIR; S35164; S35164.			
DR	PDB; 1BI5; X-ray; A=1-389.			
DR	PDB; 1BQ6; X-ray; A=2-389.			
DR	PDB; 1CGK; X-ray; A=1-389.			
DR	PDB; 1CGZ; X-ray; A=1-389.			

DR PDB; 1CHW; X-ray; A/B=1-389.  
 DR PDB; 1CML; X-ray; A=1-389.  
 DR PDB; 1SGF; X-ray; A=1-389.  
 DR PDB; 1D6H; X-ray; A=3-389.  
 DR PDB; 1D6I; X-ray; A/B=2-389.  
 DR PDB; 1186; X-ray; A=1-389.  
 DR PDB; 1188; X-ray; A/B=1-389.  
 DR PDB; 1189; X-ray; A/B=1-389.  
 DR PDB; 118B; X-ray; A/B=1-389.  
 DR PDB; 1LWX; X-ray; A=1-389.  
 DR InterPro; IPR001099; N-C synthase.  
 DR InterPro; IPR011141; PKS III.  
 DR Pfam; PF02797; Chal sti synt C; 1.  
 DR Pfam; PF00195; Chal sti synt N; 1.  
 DR PIRSF; PIRSF000451; PKS\_III; 1.  
 DR ProDom; PD000453; N-C synthase; 1.  
 DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
 KW 3D-structure; Acyltransferase; Flavonoid biosynthesis;  
 KW Multigene family; Transferase.  
 FT ACT SITE 164 164  
 FT HELIX 4 11  
 FT STRAND 18 25  
 FT STRAND 30 32  
 FT TURN 34 35  
 FT HELIX 36 43  
 FT TURN 44 45  
 FT TURN 47 48  
 FT TURN 50 62  
 FT HELIX 67 69  
 FT STRAND 74 78  
 FT HELIX 79 79  
 FT TURN 81 84  
 FT HELIX 91 117  
 FT HELIX 121 123  
 FT STRAND 126 131  
 FT HELIX 140 148  
 FT TURN 149 149  
 FT TURN 152 153  
 FT STRAND 155 160  
 FT TURN 166 179  
 FT HELIX 180 180  
 FT TURN 182 183  
 FT STRAND 185 192  
 FT HELIX 194 196  
 FT TURN 197 197  
 FT TURN 203 204  
 FT HELIX 206 214  
 FT STRAND 215 215  
 FT STRAND 218 225  
 FT TURN 230 232  
 FT STRAND 237 246  
 FT TURN 248 249  
 FT TURN 251 252  
 FT STRAND 254 255  
 FT STRAND 258 259  
 FT TURN 260 261  
 FT STRAND 262 263  
 FT STRAND 266 267  
 FT HELIX 271 287  
 FT HELIX 288 290  
 FT TURN 291 291  
 FT TURN 295 296  
 FT STRAND 299 302  
 FT HELIX 307 317  
 FT TURN 318 318  
 FT HELIX 321 324  
 FT HELIX 325 334  
 FT STRAND 336 336  
 FT HELIX 338 340  
 FT HELIX 341 355  
 FT TURN 356 357  
 FT STRAND 360 360

FT TURN 361 364  
 FT STRAND 367 374  
 FT TURN 375 377  
 FT STRAND 378 386  
 FT TURN 388 389  
 SQ SEQUENCE 389 AA; 42706 MW; E03422EB332A5408 CRC64;  
 Query Match 99.8%; Score 1994; DB 1; Length 389;  
 Best Local Similarity 99.7%; Pred. No. 7.7e-136;  
 Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKQRAEGPATIIAIGTANPANCVEQSTYDFYFKITNSHKTELKPKFORMC 60  
 DB 1 MVSVEIRKQRAEGPATIIAIGTANPANCVEQSTYDFYFKITNSHKTELKPKFORMC 60  
 QY 61 DSKMIKRYMYLITEIILKENPNVCEYMAPSLDADQMVVVEPRLGKEAAVKAKEWGQP 120  
 DB 61 DSKMIKRYMYLITEIILKENPNVCEYMAPSLDADQMVVVEPRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVYKRYMYQOQGFAGGTVLRKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVYKRYMYQOQGFAGGTVLRKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAVTFRGPSDTHLSDSLVQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSDTHLSDSLVQALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 QY 241 WTAOTIAPDSEGAIDGHLREAGLTFFHLLKOVPGIVSKNITKALVEAPEPLGISDYSNIFW 300  
 DB 241 WTAOTIAPDSEGAIDGHLREAGLTFFHLLKOVPGIVSKNITKALVEAPEPLGISDYSNIFW 300  
 QY 301 IAHPGGPAILDQVEQKLALKEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKLALKEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGLVRFPGFGPLTIETVLRSAI 389  
 DB 361 TGEGLWGLVRFPGFGPLTIETVLRSAI 389  
 RESULT 2  
 CHS4 MEDSA STANDARD; PRT; 389 AA.  
 ID CHS4 MEDSA  
 AC P30075;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 4 (EC 2.3.1.74) (Naringenin-chalcone synthase 4)  
 GN Name=CHS4;  
 OS Medicago sativa (Alfalfa).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 OX NCBI\_TaxID=3879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Iroquois; TISSUE=Root;  
 RX MEDLINE=94250839; PubMed=8193301;  
 RA McKnann H.I., Hirsch A.M.;  
 RT "Isolation of chalcone synthase and chalcone isomerase cDNAs from  
 alfalfa (Medicago sativa L.): highest transcript levels occur in young  
 roots and root tips.";  
 RL Plant Mol. Biol. 24:767-777(1994).  
 RN [2]  
 RP SEQUENCE OF 7-389 FROM N.A.  
 RX MEDLINE=93283629; PubMed=8507827;  
 RA Jungmans H., Dalkin K., Dixon R.A.;  
 RT "Stress responses in alfalfa (Medicago sativa L.). 15.  
 RT Characterization and expression patterns of members of a subset of the  
 chalcone synthase multigene family.";  
 RL Plant Mol. Biol. 22:239-253(1993).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-

CC	tetrahydroxycalcone (also termed naringenin-chalcone or chalcone) which can under specific conditions spontaneously isomerize into naringenin.
CC	-I- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).
CC	-I- PATHWAY: Part of the biosynthetic pathway for all classes of flavonoids, a large class of secondary plant metabolites, many of which are brightly colored.
CC	-I- DEVELOPMENTAL STAGE: Highest expression in young root tips.
CC	-I- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; U01021; AAB41559.1; -
DR	EMBL; L02903; AAA02825.1; ALT_INTT.
DR	PIR; S44370; S44370.
DR	HSSP; P30074; 1CGZ.
DR	InterPro; IPR001099; N-C synthase.
DR	InterPro; IPR011141; PKS_III.
DR	Pfam; PF02797; Chal_sti_synt_C; 1.
DR	Pfam; PF00195; Chal_sti_synt_N; 1.
DR	PIRSF; PIRSF000451; PKS_III; 1.
DR	ProDom; PD000453; N-C synthase; 1.
DR	PROSITE; PS00443; CHALCONE_SYNTH; 1.
DR	Acyltransferase; Flavonoid biosynthesis; Multigene family;
KW	Transferase.
FT	ACT SITE 164 164 By similarity.
FT	ACT SITE 389 AA; 42735 MW; A969APD7D7483C49 CRC64;
FT	SEQUENCE 389 AA; 42735 MW; A969APD7D7483C49 CRC64;
Qy	Query Match 98.6%; Score 1969; DB 1; Length 389;
Db	Best Local Similarity 98.5%; Pred. No. 4.9e-134;
Qy	Matches 383; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db	1 MVSVEIRKQARAGPATILAIQTANPANCVQEQTYPDPFYFKIINSEHKTELKEKFORMC 60
Qy	1 MVSVEIRKQARAGPATILAIQTANPANCVQEQTYPDPFYFKIINSEHKTELKEKFORMC 60
Db	61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPLGKEAAVKAKEWGQP 120
Qy	61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPLGKEAAVKAKEWGQP 120
Db	121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYKRYMYQOQGFAGTGLRLAKDLAEN 180
Qy	121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYKRYMYQOQGFAGTGLRLAKDLAEN 180
Db	121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYKRYMYQOQGFAGTGLRLAKDLAEN 180
Qy	181 NKGARVLVVCSEVTAATFRGSDTHLSLQALFGDGAALIVGSDVPVEIEKPIFEMV 240
Db	181 NKGARVLVVCSEVTAATFRGSDTHLSLQALFGDGAALIVGSDVPVEIEKPIFEMV 240
Qy	241 WTAQTIAPDSGAIIDGHLREAGLTFHLLKOVPGIVSKNITKALVEAPEPLGISDYSIFW 300
Db	241 WTAQTIAPDSGAIIDGHLREAGLTFHLLKOVPGIVSKNIDKALVEAQPLGISDYSIFW 300
Qy	301 IAHFGGPAILDQVQKALKPEKKNATREVLSEYGNMSSACVLFIIDEMRKKTQNGLKT 360
Db	301 IAHFGGPAILDQVQKALKPEKKNATREVLSEYGNMSSACVLFIIDEMRKKTQNGLKT 360
Qy	361 TGEGLWGVLFPGFGPLTIETVILRSVAI 389
Db	361 TGEGLWGVLFPGFGPLTIETVILRSVAI 389
RESULT 3	
Q9FEW4	PRELIMINARY; PRT; 389 AA.
ID Q9FEW4	
AC Q9FEW4	
DT 01-MAR-2001 (TrEMBLrel. 16, Created)	

GN Name=CHS8;  
OS Medicago sativa (Alfalfa).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93283629; PubMed=8507827;  
RA Jungmans H., Dalkin K., Dixon R.A.;  
RT "Stress responses in alfalfa (Medicago sativa L.). 15.  
RT Characterization and expression patterns of members of a subset of the  
RT chalcone synthase multigene family.";  
RL Plant Mol. Biol. 22:239-253(1993).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydroxylchalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J02904; AAA02826.1; -;  
DR PIR; S35166; S35166.  
DR HSSP; P30074; 1CGZ.  
DR InterPro; IPR01099; N-C synthase.  
DR InterPro; IPR011141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF000451; PKS\_III; 1.  
DR ProDom; PD000453; N-C synthase; 1.  
DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
KW Transferase. 164 164 By similarity.  
FT ACT\_SITE 389 AA; 42549 MW; ASDFC41B4A9DBF2C CRC64;  
SQ SEQUENCE 389 AA; 42549 MW; ASDFC41B4A9DBF2C CRC64;  
  
Query Match 97.2%; Score 1941; DB 1; Length 389;  
Best Local Similarity 97.4%; Pred. No. 5.2e-132;  
Matches 379; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
  
QY 1 MVSVSIRKQAEAGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKKEPQRC 60  
DB 1 MVSVSIRKQAEAGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKKEPQRC 60  
  
QY 61 DKSMIKRRWYLTETILKENPNVCEYMAPSLDARDQMVVVEVPRGLCKEAAVKAIEWGQP 120  
DB 61 DKSMIKRRWYLTETILKENPNVCEYMAPSLDARDQMVVVEVPRGLCKEAAVKAIEWGQP 120  
  
QY 121 KSKITHLIVCTTSGVDMPCADYQTLTKLGLRPPVKKRYMMYQQGXFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPCADYQTLTKLGLRPPVKKRYMMYQQGXFAGGTVLRLAKDLAEN 180  
  
QY 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVTA VTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
  
QY 241 WTAQTTAPSEGNIDGHLREAGLTFTLLKDVPGIVSKNTTKALVEAFELGHSYNSIFW 300  
DB 241 WTAQTTAPSEGNIDGHLREAGLTFTLLKDVPGIVSKNTTKALVEAFELGHSYNSIFW 300  
  
QY 301 IAHGGGPAILLDQVEQKALKPKKMKATREVLSEYGNMSSACVLVILDEMRRKKSADGLKT 360

Db 301 IAHGGGPAILLDQVEQKALKPKKMKATREVLSEYGNMSSACVLVILDEMRRKKSADGLKT 360  
QY 361 TGEGLGVGLFGPFGPGLTIETVLRSAI 389  
DB 361 TGEGLGVGLFGPFGPGLTIETVLRSAI 389  
  
RESULT 5  
CHS5\_MEDSA STANDARD; PRT; 389 AA.  
AC PS1078;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 4-2 (EC 2.3.1.74) (Naringenin-chalcone synthase 4-2).  
DE 2).  
GN Name=CHS4-2; Synonyms=CHS1;  
OS Medicago sativa (Alfalfa).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
OX NCBI\_TaxID=3879;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nagvazenas; TISSUE=Leaf;  
RX MEDLINE=94100583; PubMed=8274775;  
RA Esnault R., Buffard D., Breda C., Sallaud C., El-Turk J.,  
RA Kondorosi A.;  
RT "Pathological and molecular characterizations of alfalfa interactions  
RT with compatible and incompatible bacteria, Xanthomonas campestris pv.  
RT alfalfae and Pseudomonas syringae pv. pisi.";  
RL Mol. Plant Microbe Interact. 6:655-664(1993).  
RN [2]  
RP SEQUENCE OF 19-389 FROM N.A.  
RC STRAIN=cv. Iroquois; TISSUE=Root nodules;  
RX MEDLINE=94250839; PubMed=8193301;  
RA McKhann H.I., Hirsch A.M.;  
RT "Isolation of chalcone synthase and chalcone isomerase cDNAs from  
RT alfalfa (Medicago sativa L.): highest transcript levels occur in young  
RT roots and root tips.";  
RL Plant Mol. Biol. 24:767-777(1994).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydroxylchalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- DEVELOPMENTAL STAGE: Highest expression in young root tips.  
CC -!- INDUCTION: By infection.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X68106; CAA48226.1; -;  
DR EMBL; J01019; AAB41560.1; -;  
DR PIR; S26414; S26414.  
DR HSSP; P30074; 1BI5.  
DR InterPro; IPR01099; N-C synthase.  
DR InterPro; IPR011141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF000451; PKS\_III; 1.  
DR ProDom; PD000453; N-C synthase; 1.

DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
 KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 FT ACT SITE 164 By similarity.  
 FT CONFLICT 86 I -> Y (in Ref. 2).  
 FT CONFLICT 94 W -> R (in Ref. 2).  
 SQ SEQUENCE 389 AA; 42661 MW; B33674EC7709F170 CRC64;

Query Match 97.0%; Score 1938; DB 1; Length 389;  
 Best Local Similarity 97.2%; Pred. No. 8.6e-132;  
 Matches 378; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKQFQRC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKQFQRC 60  
 QY 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQMDVVVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQMDVVVPRLGKEAAVKAKEWGOP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLCVSEVTAATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 DB 181 NKGARVLCVSEVTAATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGVLFPGFGLTIETVTLRSVAI 389  
 DB 361 TGEGLWGVLFPGFGLTIETVTLRSVAI 389

RESULT 6  
 CHS5\_PEA  
 ID CHS5\_PEA STANDARD; PRT; 389 AA.  
 AC Q23884;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 5 (EC 2.3.1.74) (Naregenin-chalcone synthase 5).  
 GN Names=CHS5;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=cv. Midoriisui;  
 RX MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;  
 RA Ito M., Ichinose Y., Kato H., Shiraiishi T., Yamada T.;  
 RT "Molecular evolution and functional relevance of the chalcone synthase genes of pea."  
 RL Mol. Gen. Genet. 255:28-37(1997).  
 CC -1- FUNCTION: The primary product of this enzyme is 4,2',4',6'-tetrahydroxychalcone (also termed naringenin-chalcone or chalcone) which can under specific conditions spontaneously isomerize into naringenin.  
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).  
 CC -1- PATHWAY: Part of the biosynthetic pathway for all classes of flavonoids, a large class of secondary plant metabolites, many of which are brightly colored.  
 CC -1- SIMILARITY: Belongs to the chalcone/stilbene synthases family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; D88262; BAA22044.1; -.  
 DR HSSP; P30074; ICGZ.  
 DR InterPro; IPR001099; N-C\_synthase.  
 DR InterPro; IPR011141; PKS\_III.  
 DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 DR PIRSF; PIRSF000451; PKS\_III; 1.  
 DR PRODOM; PD000453; N-C\_synthase; 1.  
 DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
 KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
 FT ACT SITE 164 164 By similarity.  
 SQ SEQUENCE 389 AA; 42816 MW; 32272DC9B45BDEA CRC64;

Query Match 97.0%; Score 1937; DB 1; Length 389;  
 Best Local Similarity 96.4%; Pred. No. 1e-131;  
 Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKQFQRC 60  
 DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKQFQRC 60  
 QY 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQMDVVVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLITEELKENPNVCYMAPSLDARDQMDVVVPRLGKEAAVKAKEWGOP 120  
 QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPPVYKRYMYQQXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLCVSEVTAATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 DB 181 NKGARVLCVSEVTAATFRGSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPGIVSKNITKALVEAFEPGLIGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPAIVSKNIDKALVEAFQPLIGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGVLFPGFGLTIETVTLRSVAI 389  
 DB 361 TGEGLWGVLFPGFGLTIETVTLRSVAI 389

RESULT 7  
 CHS2\_PEA  
 ID CHS2\_PEA STANDARD; PRT; 389 AA.  
 AC Q01287;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 2 (EC 2.3.1.74) (Naregenin-chalcone synthase 2).  
 GN Name=CHS2;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=cv. Midoriisui; TISSUE=Epicotyl;

```

RX MEDLINE=92256801; PubMed=1581561;
RA Ichinose Y., Kawanata S., Yamada T., An C., Kajiwa T., Shiraishi T.,
RA Oku H.;
RA "Molecular cloning of chalcone synthase cDNAs from Pisum sativum.";
RL Plant Mol. Biol. 18:1009-1012(1992).
RN [2]
RP SEQUENCE OF 1-128 FROM N.A.
RC STRAIN=cv. Midoriisui; TISSUE=Epicotyl;
RX MEDLINE=93222478; PubMed=8467077;
RA An C., Ichinose Y., Yamada T., Tanaka Y., Shiraishi T., Oku H.;
RT "Organization of the genes encoding chalcone synthase in Pisum
RT sativum.";
RL Plant Mol. Biol. 21:789-803(1993).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63334; CAA44934.1; -.
DR EMBL; D10662; BAA01513.1; -.
DR PIR; S20932; S20932.
DR HSSP; P30074; 1CGZ.
DR InterPro; IPR001099; N-C synthase.
DR InterPro; IPR011141; PKS_III.
DR Pfam; PF02797; Chal_sti_synt_C; 1.
DR Pfam; PF00195; Chal_sti_synt_N; 1.
DR PIRSF; PIRSF000451; PKS_III; 1.
DR ProDom; PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
FT ACT_SITE 164 164 By similarity.
FT ACT_SITE 389 AA; 42844 MW; 0585EC37CA4F072 CRC64;
SQ SEQUENCE 389 AA; 42844 MW; 0585EC37CA4F072 CRC64;

Query Match 96.7%; Score 1932; DB 1; Length 389;
Best Local Similarity 96.4%; Pred. No. 2.3e-131;
Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKQFQRC 60
DB 1 MVTVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKQFQRC 60
QY 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120
DB 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMMYQQCGFAGGTVRLAKDLAEN 180
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPVYKRYMMYQQCGFAGGTVRLAKDLAEN 180
QY 181 NKGARVLVVCSEVAVTFRGPSTHLDLSLVQALFGDGAALIVGSDPVEIEKPIFENV 240
DB 181 NKNARVLVVCSEVAVTFRGPSTHLDLSLVQALFGDGAALIVGSDPVEIEKPIFENV 240
QY 241 WTAQTAPDSEGAIDGHLREAGITFHLKDVPGIVSKNITKALVEAFELGHSYNSIFW 300
DB 241 WTAQTAPDSEGAIDGHLREAGITFHLKDVPGIVSKNIDKALVEAFELGHSYNSIFW 300
QY 301 IAHFGGPAILDQVEQKALKPEKONATREVLSEYGNMSSACVLFILDEMRKKSQTONGLKT 360

```

```

DB 301 IAHFGGPAILDQVEQKALKPEKMRATREVLSEYGNMSSACVLFILDEMRKKSQTONGLKT 360
QY 361 TGEGLWGVLFPGFGPLTIETVTLRSVAI 389
DB 361 TGEGLWGVLFPGFGPLTIETVTLRSVAI 389

RESULT 8
CHS3_PEA CHS3_PEA STANDARD; PRT; 389 AA.
ID CHS3_PEA AC O23883;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chalcone synthase 3 (EC 2.3.1.74) (Naregenin-chalcone synthase 3).
GN Name=CHS3;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Midoriisui;
RX MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;
RA Ito M., Ichinose Y., Kato H., Shiraishi T., Yamada T.;
RT "Molecular evolution and functional relevance of the chalcone synthase
RT genes of pea.";
RL Mol. Gen. Genet. 255:28-37(1997).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D88261; BAA22043.1; -.
DR HSSP; P30074; 1BI5.
DR InterPro; IPR001099; N-C synthase.
DR InterPro; IPR011141; PKS_III.
DR Pfam; PF02797; Chal_sti_synt_C; 1.
DR Pfam; PF00195; Chal_sti_synt_N; 1.
DR PIRSF; PIRSF000451; PKS_III; 1.
DR ProDom; PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
KW Transferase.
FT ACT_SITE 164 164 By similarity.
FT ACT_SITE 389 AA; 42802 MW; 70720C8BCC15CAB8 CRC64;
SQ SEQUENCE 389 AA; 42802 MW; 70720C8BCC15CAB8 CRC64;

Query Match 96.6%; Score 1930; DB 1; Length 389;
Best Local Similarity 96.1%; Pred. No. 3.3e-131;
Matches 374; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKQFQRC 60
DB 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKQFQRC 60
QY 61 DKSMIKRRYMYLTEETILKENPNVCYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGOP 120

```



Db 61 DKSMINRRMYLTEELKENPVCSEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPYVRYKRYMYQQXFAAGGTVLRKLAKDAEN 180  
 Db 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPYVRYKRYMYQQXFAAGGTVLRKLAKDAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDPVPTEKEPIFEMV 240  
 Db 181 NKGARVLVVCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDPVPTEKEPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPAIVSKNIDKALVEAFEPGLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPEKKNATREVLSEYGNMSACVLFILDEMRKKSQTQGLTKT 360  
 Db 301 IAHPGGPAILDQVEQKALPEKKNATREVLSEYGNMSACVLFILDEMRKKSQTQGLTKT 360  
 QY 361 TGEGLWGVLFPGFGPLTETVTLRSVAI 389  
 Db 361 TGEGLWGVLFPGFGPLTETVTLRSVAI 389

## RESULT 9

CHS2\_TRISU STANDARD; PRT; 389 AA.  
 ID CHS2\_TRISU STANDARD; PRT; 389 AA.  
 AC P51084;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).  
 GN Names:CHS2;  
 OS Trifolium subterraneum (Subterranean clover).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.  
 OX NCBI\_TaxID=3900;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Karridale; TISSUE=Leaf, and Stem;  
 RX MEDLINE=94171050; PubMed=8125320; DOI=10.1016/0378-1119(94)90785-4;  
 RA Arioli T., Howles P.A., Weisman J.J., Rolfe B.G.;  
 RT "In Trifolium subterraneum, chalcone synthase is encoded by a  
 RT multigene family";  
 RL Gene 138:79-86(1994).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
 CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).  
 CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M91194; AAA18177.1; -;  
 CC HSP; P30074; IBI5.  
 CC InterPro; IPR01039; N-C synthase.  
 CC InterPro; IPR01141; PKS III.  
 CC Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
 CC Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
 CC FIRSF; FIRSF00451; PKS\_III; 1.  
 CC ProDom; PD000453; N-C synthase; 1.  
 CC PROSITE; PS00441; CHALCONE\_SYNTH; 1.

KW Acyltransferase; Flavanoid biosynthesis; Multigene family;  
 KW Transferase.  
 FT ACT SITE 164 164 By similarity.  
 SQ SEQUENCE 389 AA; 42728 MW; 51200DE3B9C58404 CRC64;  
 Query Match 96.5%; Score 1927; DB 1; Length 389;  
 Best Local Similarity 96.7%; Pred. No. 5.4e-131;  
 Matches 376; Conservative 3; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAQRAEGPATILGATNPANCVQSTQSTVDFYFKITITSEHKTELKEKEQRM 60  
 Db 1 MVSVEIRKAQRAEGPATILGATNPANRVQATYDFYFKITITSEHKTELKEKEQRM 60  
 QY 61 DKSMIKRRMYLTEELKENPVCSEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 Db 61 DKSMIKRRMYLTEELKENPVCSEYMAPSLDARQDMVVVPRRLGKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPYVRYKRYMYQQXFAAGGTVLRKLAKDAEN 180  
 Db 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRLPYVRYKRYMYQQXFAAGGTVLRKLAKDAEN 180  
 QY 181 NKGARVLVVCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDPVPTEKEPIFEMV 240  
 Db 181 NKGARVLVVCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDPVPTEKEPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 Db 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNIDKALVEAFEPGLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPEKKNATREVLSEYGNMSACVLFILDEMRKKSQTQGLTKT 360  
 Db 301 IAHPGGPAILDQVEQKALPEKKNATREVLSEYGNMSACVLFILDEMRKKSQTQGLTKT 360  
 QY 361 TGEGLWGVLFPGFGPLTETVTLRSVAI 389  
 Db 361 TGEGLWGVLFPGFGPLTETVTLRSVAI 389  
 RESULT 10  
 CHS1\_PEA  
 ID CHS1\_PEA STANDARD; PRT; 389 AA.  
 AC Q01286;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).  
 GN Name:CHS1;  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosid I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Midoriisui; TISSUE=Epicotyl;  
 RX MEDLINE=92256801; PubMed=1581561;  
 RA Ichinose Y., Kawamata S., Yamada T., An C., Kajiura T., Shiraishi T.,  
 RA Oku H.;  
 RT "Molecular cloning of chalcone synthase cDNAs from Pisum sativum";  
 RL Plant Mol. Biol. 18:1009-1012(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93222478; PubMed=8467077;  
 RA An C., Ichinose Y., Yamada T., Tanaka Y., Shiraishi T., Oku H.;  
 RT "Organization of the genes encoding chalcone synthase in Pisum  
 RT sativum";  
 RL Plant Mol. Biol. 21:789-803(1993).  
 CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
 CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)  
 CC which can under specific conditions spontaneously isomerize into  
 CC naringenin.  
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
 CC naringenin-chalcone + 3 CO(2).

CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
 CC flavonoids, a large class of secondary plant metabolites, many of  
 CC which are brightly colored.  
 CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; X63333; CAA44933.1; -;  
 DR EMBL; D10661; BAA01512.1; -;  
 DR PIR; S33610; S33610.  
 DR HSP; P30074; IB15.

DR InterPro; IPR001099; N-C synthase.

DR InterPro; IPR011141; PKS\_III.

DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.

DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.

DR PIRSF; PIRSF000451; PKS\_III; 1.

DR ProDom; PD000453; N-C synthase; 1.

DR PROSITE; PS00441; CHALCONE\_SYNTH; 1.

KW Acyltransferase; Flavonoid biosynthesis; Multigene family;

FT ACT SITE 164 164 By similarity.

SEQUENCE 389 AA; 42802 MW; BEE7FD08B0AEB4 CRC64;

Query Match 96.4%; Score 1926; DB 1; Length 389;

Best Local Similarity 96.1%; Pred. No. 6.3e-131;

Matches 374; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVSVEIRKAORAEAGPATILTAIGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 DB 1 MVSVEIRKAORAEAGPATILTAIGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60

QY 61 DKSIMKRRYMYLTHEILKENPNCVEYMAPSLDARDQMDVMVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSIMKRRYMYLTHEILKENPNCVEYMAPSLDARDQMDVMVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQOGXCFAGGTGVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQOGXCFAGGTGVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFENV 240  
 DB 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360

QY 361 TGELEWGLVFGFGPLGTETVLRSAI 389  
 DB 361 TGELEWGLVFGFGPLGTETVLRSAI 389

RESULT 11

ID CHS9 MEDSA

AC P30077; STANDARD; PRT; 389 AA.

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Chalcone synthase 9 (EC 2.3.1.74) (Naringenin-chalcone synthase 9).

GN Name:CHS9;

OS Medicago sativa (Alfalfa).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eucosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 ON NCBI\_TaxID=3879;

OX RN

RP SEQUENCE FROM N.A.

RX MEDLINE=93283629; PubMed=8507827;

RA Junghans H., Dalkin K., Dixon R.A.;

RT "Stress responses in alfalfa (Medicago sativa L.). 15.

RT Characterization and expression patterns of members of a subset of the

RT chalcone synthase multigene family.";

RL Plant Mol. Biol. 22:239-253(1993).

CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-

CC tetrahydroxychalcone (also termed naringenin-chalcone or chalcone)

CC which can under specific conditions spontaneously isomerize into

CC naringenin.

CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +

CC naringenin-chalcone + 3 CO(2).

CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of

CC flavonoids, a large class of secondary plant metabolites, many of

CC which are brightly colored.

CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L02905; AAA02827.1; -;

DR PIR; S35167; S35167.

DR HSP; P30074; IB15.

DR InterPro; IPR001099; N-C synthase.

DR InterPro; IPR011141; PKS\_III.

DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.

DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.

DR PIRSF; PIRSF000451; PKS\_III; 1.

DR ProDom; PD000453; N-C synthase; 1.

DR PROSITE; PS00441; CHALCONE\_SYNTH; 1.

KW Acyltransferase; Flavonoid biosynthesis; Multigene family;

FT ACT SITE 164 164 By similarity.

SEQUENCE 389 AA; 42721 MW; BB461B53B8C63002 CRC64;

Query Match 96.4%; Score 1926; DB 1; Length 389;

Best Local Similarity 95.9%; Pred. No. 6.3e-131;

Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVSVEIRKAORAEAGPATILTAIGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
 DB 1 MVSVEIRKAORAEAGPATILTAIGTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60

QY 61 DKSIMKRRYMYLTHEILKENPNCVEYMAPSLDARDQMDVMVVEVPRLGKEAAVKAKEWGQP 120  
 DB 61 DKSIMKRRYMYLTHEILKENPNCVEYMAPSLDARDQMDVMVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQOGXCFAGGTGVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQOGXCFAGGTGVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFENV 240  
 DB 181 NKGARVLVVCSEVTATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLKDVPGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKKATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360

```

QY 361 TGGLEWGLVFGFGPGLTIETVLRVAI 389
DB 361 TGGLEWGLVFGFGPGLTIETVLRVAI 389

RESULT 12
CHS4_PEA
ID CHS4_PEA STANDARD; PRT; 389 AA.
AC O23882;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Chalcone synthase 4 (EC 2.3.1.74) (Naringenin-chalcone synthase 4).
GN NamesCHS4;
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3886;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Midoriisui;
RC MEDLINE=97374447; PubMed=9230896; DOI=10.1007/s004380050471;
RA Ito M., Ichinose Y., Kato H., Shiraiishi T., Yamada T.;
RT "Molecular evolution and functional relevance of the chalcone synthase
RT genes of pea.";
RL Mol. Gen. Genet. 255:28-37(1997).
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)
CC which can under specific conditions spontaneously isomerize into
CC naringenin.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of
CC flavonoids, a large class of secondary plant metabolites, many of
CC which are brightly colored.
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D88260; BAA22042.1; --
CC HSSP; P30074; 1B15.
CC InterPro; IPR001099; N-C_synthase.
CC Pfam; PF02797; Chal_sti_synth_C; 1.
CC PIRSF; PIRSF000453; PKS_III; 1.
CC ProDom; PD000453; N-C_synthase; 1.
CC PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;
FT ACT SITE 164 164 By similarity.
SQ SEQUENCE 389 AA; 42856 MW; 7DC9D3F266CD125 CRC64;

Query Match 96.4%; Score 1925; DB 1; Length 389;
Best Local Similarity 95.9%; Pred. No. 7.5e-131;
Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 1 MVSVSIRKQRAEGPATILAIATNPANCVQSTYDPDFYFKITSEHKTELKEKQRM 60
DB 1 MVSVSIRKQRAEGPATILAIATNPANCVQSTYDPDFYFKITSEHKTELKEKQRM 60
QY 61 DKSMIKRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAIKWGQP 120
DB 61 DKSMIKRRYMYLTELKENPNVCYMAPSLDARQDMVVVEVPRLGEAAVKAIKWGQP 120
QY 121 KSKITHLVCTTSGVDMPGADYQLTKLLGLRPPYVKRYMYQQGCFAGGTVLRKDLAEN 180
DB 121 KSKITHLVCTTSGVDMPGADYQLTKLLGLRPPYVKRYMYQQGCFAGGTVLRKDLAEN 180

```

FT ACT\_SITE 164 164 By similarity.  
SQ SEQUENCE 389 AA; 42672 MW; A67A9155A89E68E4 CRC64;  
Query Match 96.1%; Score 1920; DB 1; Length 389;  
Best Local Similarity 95.9%; Pred. No. 1.7e-130;  
Matches 373; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVAEIRKAQRAEGPATILAI GTANPANKVEQATYDPDFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGP 120  
DB 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGP 120  
QY 121 KSKITHLIVCTTSGVDMPCADYQTLGLGLRPVVKRYMYQGXFGAGTVLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPCADYQTLGLGLRPVVKRYMYQGXFGAGTVLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVTAVTFRGPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNIDKALVEAFQPLNISYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389

RESULT 14  
CHS6\_PEA  
ID CHS6\_PEA STANDARD; PRT; 389 AA.  
AC Q01288;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 6 (EC 2.3.1.74) (Naringenin-chalcone synthase 6).  
GN Names=CHS6;  
OS Pisum sativum (Garden pea).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
OX NCBI\_TaxID=3888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Midoriusul; TISSUE=epicotyl;  
RX MEDLINE=92256801; PubMed=1581561;  
RA Ichinose Y., Kawamata S., Yamada T., An C., Kajiura T., Shiraishi T.,  
RA Oku H.;  
RT "Molecular cloning of chalcone synthase cDNAs from Pisum sativum.";  
RL Plant Mol. Biol. 18:1009-1012(1992).  
CC -!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-  
CC tetrahydrochalcone (also termed naringenin-chalcone or chalcone)  
CC which can under specific conditions spontaneously isomerize into  
CC naringenin.  
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +  
CC naringenin-chalcone + 3 CO(2).  
CC -!- PATHWAY: Part of the biosynthetic pathway for all classes of  
CC flavonoids, a large class of secondary plant metabolites, many of  
CC which are brightly colored.  
CC -!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; X63335; CAA44935.1; -;  
CC PIR; S20933; S20933.  
CC DR HSSP; P30074; 1CGZ.  
DR InterPro; IPR001099; N-C synthase.  
DR InterPro; IPR011141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF000451; PKS\_III; 1.  
DR ProDom; PD000453; N-C synthase; 1.  
DR PROSITE; PS00441; CHALCONE SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
KW Transferase.  
KW ACT\_SITE 164 164 By similarity.  
SQ SEQUENCE 389 AA; 42761 MW; 73FC26AFDF82EEDD CRC64;  
Query Match 95.9%; Score 1915; DB 1; Length 389;  
Best Local Similarity 95.4%; Pred. No. 4e-130;  
Matches 371; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAI GTATPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGP 120  
DB 61 DKSMIKRMYLTYTEILKENPNVCEYMAPSLDARDQMVVVEPRLGKEAAVKAKEWGP 120  
QY 121 KSKITHLIVCTTSGVDMPCADYQTLGLGLRPVVKRYMYQGXFGAGTVLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPCADYQTLGLGLRPVVKRYMYQGXFGAGTVLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTAVTFRGPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVTAVTFRGPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNINKALVEAFQPLNIDYNSIFW 300  
QY 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389

RESULT 15  
CHS5\_TRISU  
ID CHS5\_TRISU STANDARD; PRT; 389 AA.  
AC P51087;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Chalcone synthase 5 (EC 2.3.1.74) (Naringenin-chalcone synthase 5).  
GN Name=CHS5;  
OS Trifolium subterraneum (Subterranean clover).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.  
OX NCBI\_TaxID=3900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Kariadale; TISSUE=Leaf, and Stem;  
RX MEDLINE=95232194; PubMed=7716240; DOI=10.1104/pp.107.3.1035;  
RA Howles P.A., Arioli T., Weinman J.J.;  
RT "Nucleotide sequence of additional members of the gene family encoding  
RT chalcone synthase in Trifolium subterraneum.";

Plant Physiol. 107:1035-1036(1995)).  
-!- FUNCTION: The primary product of this enzyme is 4,2',4',6'-tetrahydrochalcone (also termed naringenin-chalcone or chalcone) which can under specific conditions spontaneously isomerize into naringenin.  
-!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).  
-!- PATHWAY: Part of the biosynthetic pathway for all classes of flavonoids, a large class of secondary plant metabolites, many of which are brightly colored.  
-!- INDUCTION: By wounding and Rhizobium infection.  
-!- SIMILARITY: Belongs to the chalcone/stilbene synthases family.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; L24517; AA073939.1; -.  
DR HSSP; P30074; 1B15  
DR InterPro; IPR01099; N-C\_synthase.  
DR InterPro; IPR01141; PKS\_III.  
DR Pfam; PF02797; Chal\_sti\_synt\_C; 1.  
DR Pfam; PF00195; Chal\_sti\_synt\_N; 1.  
DR PIRSF; PIRSF00451; PKS\_III; 1.  
DR PRODOM; PD000453; N-C\_synthase; 1.  
DR PROSITE; PS00441; CHALCONE\_SYNTH; 1.  
KW Acyltransferase; Flavonoid biosynthesis; Multigene family;  
FT ACT SITE 164 164 By similarity.  
SQ SEQUENCE 389 AA; 42656 MW; EEAC4AA9717D4CA9 CRC64;  
  
Query Match 95.8%; Score 1914; DB 1; Length 389;  
Best Local Similarity 95.6%; Pred. No. 4.7e-130;  
Matches 372; Conservative 7; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 MVSVAIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVAIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
  
QY 61 DKSMIKRRYMYLTTEEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLTTEEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQP 120  
  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMMYQQXFPAGGTVLRKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMMYQQXFPAGGTVLRKDLAEN 180  
  
QY 181 NKGARVLVWCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDVPVEIEKPIFEMV 240  
DB 181 NKGARVLVWCSEVTATFRGPDTHLDSLVGQALFGDGAALIVGSDVPVEIEKPIFEMV 240  
  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGIDSYNISIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGIDSYNISIFW 300  
  
QY 301 IAHPGGPAILDQVEQKALKPEKONATREVLSEYGNMSACVLFILDEMRRKKSSTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKPEKONATREVLSEYGNMSACVLFILDEMRRKKSSTONGLKT 360  
  
QY 361 TGEGLDWGLVFGFGPLTIETVTVLSVAI 389  
DB 361 TGEGLDWGLVFGFGPLTIETVTVLSVAI 389

Search completed: April 22, 2005, 21:44:16  
Job time : 174 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:35:47 ; Search time 42 Seconds  
(without alignments)  
691.393 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVSIRKQRAEGPATIL.....LFGFGELTETVLRSAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCFUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1508.5	75.5	394	US-09-252-816A-1	Sequence 1, Appli
2	1508.5	75.5	394	US-09-837-654-1	Sequence 1, Appli
3	477	23.9	151	US-08-321-358-2	Sequence 2, Appli
4	314	15.7	349	US-08-994-035C-5	Sequence 5, Appli
5	314	15.7	349	US-09-395-861-5	Sequence 5, Appli
6	310	15.5	349	US-08-494-907-12	Sequence 12, Appli
7	310	15.5	349	PCT-US96-10986-12	Sequence 12, Appli
8	304	15.2	360	US-09-902-540-13778	Sequence 20, Appli
9	183	9.2	506	US-09-877-476-20	Sequence 20, Appli
10	181	9.1	506	US-09-877-476-16	Sequence 16, Appli
11	181	9.1	506	US-09-877-476-22	Sequence 22, Appli
12	180	9.0	506	US-09-877-476-40	Sequence 40, Appli
13	179.5	9.0	521	US-08-657-749D-6	Sequence 6, Appli
14	179	9.0	506	US-09-877-476-12	Sequence 12, Appli
15	178	8.9	505	US-09-877-476-6	Sequence 6, Appli
16	177	8.9	506	US-09-877-476-8	Sequence 8, Appli
17	177	8.9	506	US-09-877-476-14	Sequence 14, Appli
18	177	8.9	506	US-09-877-476-38	Sequence 38, Appli
19	177	8.9	506	US-09-877-476-42	Sequence 42, Appli
20	176	8.8	475	US-08-657-749D-11	Sequence 11, Appli
21	176	8.8	475	US-08-657-749D-13	Sequence 13, Appli
22	176	8.8	505	US-09-877-476-18	Sequence 18, Appli
23	176	8.8	506	US-09-877-476-26	Sequence 26, Appli
24	176	8.8	506	US-09-877-476-34	Sequence 34, Appli
25	173	8.7	506	US-09-877-476-24	Sequence 24, Appli
26	173	8.7	506	US-09-877-476-32	Sequence 32, Appli
27	172	8.6	505	US-09-877-476-10	Sequence 10, Appli

28	172	8.6	506	4	US-09-877-476-4	Sequence 4, Appli
29	171	8.6	506	3	US-08-888-998-2	Sequence 2, Appli
30	171	8.6	506	3	US-09-362-633-2	Sequence 2, Appli
31	171	8.6	506	4	US-09-877-476-2	Sequence 2, Appli
32	171	8.6	506	4	US-09-877-476-28	Sequence 28, Appli
33	170	8.5	506	4	US-09-877-476-30	Sequence 30, Appli
34	169	8.5	506	4	US-09-877-476-36	Sequence 36, Appli
35	167	8.4	333	3	US-09-134-001C-3946	Sequence 3946, Ap
36	166	8.3	493	3	US-08-868-373-4	Sequence 4, Appli
37	165	8.3	524	4	US-08-657-749D-4	Sequence 4, Appli
38	164	8.2	505	4	US-08-657-749D-25	Sequence 25, Appli
39	162.5	8.1	497	3	US-09-058-947A-4	Sequence 4, Appli
40	162.5	8.1	500	3	US-08-868-373-12	Sequence 12, Appli
41	156.5	7.8	516	3	US-08-868-373-14	Sequence 14, Appli
42	156	7.8	498	4	US-08-657-749D-23	Sequence 23, Appli
43	155	7.8	313	4	US-08-970-647-2	Sequence 2, Appli
44	154.5	7.7	504	3	US-08-868-373-6	Sequence 6, Appli
45	151	7.6	313	4	US-10-138-701-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-252-816A-1  
; Sequence 1, Application US/09252816A  
; Patent No. 6265633  
; GENERAL INFORMATION:  
; APPLICANT: OKADA, Yukio  
; APPLICANT: ITO, Kazutoshi  
; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE  
; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 1959-0008-0  
; CURRENT APPLICATION NUMBER: US/09/252,816A  
; CURRENT FILING DATE: 1999-02-19  
; PRIOR APPLICATION NUMBER: JP HEI 10-37266  
; PRIOR FILING DATE: 1998-02-19  
; PRIOR APPLICATION NUMBER: JP HEI 10-174235  
; PRIOR FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Humulus lupulus  
; US-09-252-816A-1

Query Match	75.5%	Score 1508.5;	DB 3;	Length 394;			
Best Local Similarity	72.1%;	Pred. No. 5.2e-153;					
Matches 279;	Conservative 59;	Mismatches 48;	Indels 1;	Gaps 1;			
Qy	2	VSVSEIRKQRAEGPATILAI	GAITANPANCVEG	STPDFYFKITNSHKT	ELKEKFORMCD 61		
Db	4	VTVQIRKQRAEGPATILAI	GAITAVPANCFCN	QADFPDYFRTVTKSE	HTDLKKKFORMCE 63		
Qy	62	KSMTKRYMVLTEIILKEN	PVCEYMAPSIDARQ	DMVVEVPELGEAAVKA	IKENGQPK 121		
Db	64	KSTIKRYLHLTEHUKQ	NHCEYNAPSLNTRQ	DMVVEVPELGEAAVKA	IKENGQPK 123		
Qy	122	SKITHLIVCTSGVDM	PMGADYQTLTKLLGR	PPVYKRYMYQQGX	FAGGTVLR	LAKDAENN 181	
Db	124	SKITHLIVCTSGSID	PMGADYQCAKLLGLR	PSVKRMVLYGCGYAGG	KVLR	IAKDIAENN 183	
Qy	182	KGARLVVCSVTA	VTFRGSPDTHLSD	LVGQAALIVGSD	VPVEI-EKPI	FEMV 240	
Db	184	KGARLVVCSSEITACI	FRGPESEKHLDC	LVGQSLFGDG	ASSVIVGAD	PDASVGERPI	PELV 243
Qy	241	WTAQTTAPDSEGA	INDGHLREAGLT	THLLKDVPGVIS	KNITKALVEA	FEPLGIS	DNSIFW 300
Db	244	SAAQTTPLNSD	GAIGHVTEAGLT	THLLRDPVPLIS	QNIKSL	EIAFTPIG	INDMNNIFW 303
Qy	301	IAHPGGPAILD	QVEQKALKPEK	MNATREVLSE	YGNSSACV	LVLFILDEMRK	IKSTQNLGT 360

Db 304 IAHPPGPAIDDEAKLEKKEKAKASREMLSEYGNMSCASVFFIVDENRKSQSGKST 363  
QY 361 TGEGLWGLVFGPGGLTETVTVLRSV 387  
Db 364 TGDGLEWGLVFGPGGLTETVTVLHSV 390

RESULT 2  
US-09-837-654-1  
; Sequence 1, Application US/09837654  
; Patent No. 6639127  
; GENERAL INFORMATION:  
; APPLICANT: OKADA, Yukio  
; APPLICANT: ITO, Kazutoshi  
; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE  
; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE  
; TITLE OF INVENTION: SAME  
; FILE REFERENCE: 1959-0008-0  
; CURRENT APPLICATION NUMBER: US/09/837,654  
; CURRENT FILING DATE: 2001-04-19  
; EARLIER APPLICATION NUMBER: 09/252,816  
; EARLIER FILING DATE: 1999-02-19  
; EARLIER APPLICATION NUMBER: JP HEI 10-174235  
; EARLIER FILING DATE: 1998-06-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Humulus lupulus  
US-09-837-654-1

Query Match 75.5%; Score 1508.5; DB 4; Length 394;  
Best Local Similarity 72.1%; Pred. No. 5.2e-153;  
Matches 279; Conservative 59; Mismatches 48; Indels 1; Gaps 1;

QY 2 VSVSEIRKQRAEGPATILAICTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEFORMCD 61  
Db 4 VIVEQIRKQRAEGPATILAICTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEFORMCE 63  
QY 62 KSMIKRYMYLTHEILKENPNVCYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPK 121  
Db 64 KSTIKRYLHTEHLKQPHLCYNAPSLNRQDMVVEVPRLGKEAAVKAKEWGQPK 123  
QY 122 SKITHLIVCTSGVDMPGADYQTLKLLGRPPYVYKRYMYTQOQXFGAGTVLRLAKOLAENN 181  
Db 124 SKITHLIVCTSGVDMPGADYQTLKLLGRPPYVYKRYMYTQOQXFGAGTVLRLAKOLAENN 183  
QY 182 KGARVLVCSSEVTAFTFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPEI-EKPTFFMV 240  
Db 184 KGARVLVCSSEVTAFTFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPEI-EKPTFFMV 243  
QY 241 WTAQTIAPDSGAIDGHLREAGLTHLLKDVPGIVSKNITKALVFAPEPLGISDYNISIFW 300  
Db 244 SAAQTILPNSDGAIAHVTEAGLTHLLKDVPGIVSKNITKALVFAPEPLGISDYNISIFW 303  
QY 301 IAHPPGPAIDDEAKLEKKEKAKASREMLSEYGNMSCASVFFIVDENRKSQSGKST 360  
Db 304 IAHPPGPAIDDEAKLEKKEKAKASREMLSEYGNMSCASVFFIVDENRKSQSGKST 363  
QY 361 TGEGLWGLVFGPGGLTETVTVLRSV 387  
Db 364 TGDGLEWGLVFGPGGLTETVTVLHSV 390

RESULT 3  
US-08-321-358-2  
; Sequence 2, Application US/08321358  
; Patent No. 5589620  
; GENERAL INFORMATION:  
; APPLICANT: Helmut Kindl, Rudiger Hain, Hans-Jorg Reif  
; TITLE OF INVENTION: BIBENZYL SYNTHASE GENES

; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SPRUNG HORN KRAMER & WOODS  
; STREET: 660 White Plains Road  
; CITY: Tarrytown  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10591-5144  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB  
; MEDIUM TYPE: storage  
; COMPUTER: HP VECTRA  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/321,358  
; FILING DATE: 11-OCT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: P 43 34 791.6 (Germany)  
; FILING DATE: 13-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kurt G. Briscoe  
; REGISTRATION NUMBER: 33,141  
; REFERENCE/DOCKET NUMBER: Bayer 9137-KGB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (914) 332-1700  
; TELEFAX: (914) 332-1844  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 151 Amino acids  
; TYPE: Amino acid  
; TOPOLOGY: Linear  
; MOLECULE TYPE: Protein  
US-08-321-358-2

Query Match 23.9%; Score 477; DB 1; Length 151;  
Best Local Similarity 58.5%; Pred. No. 6e-43;  
Matches 86; Conservative 30; Mismatches 31; Indels 0; Gaps 0;  
QY 1 MVSVSEIRKQRAEGPATILAICTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEFORMC 60  
Db 1 MPELESIKKAPRAGDFASILAIGRANPNLIEQSAYPDFYFRVTNSEHLVDLKKFKQIC 60  
QY 61 DKSMIKRYMYLTHEILKENPNVCYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPK 120  
Db 61 EKTAIRKHFVWNEBFLTANPCFSTFMDKSLNVRQEVASIBTPKLGAKAATKAIEDWGQPK 120  
QY 121 KSKITHLIVCTSGVDMPGADYQTLTKL 147  
Db 121 KSRITHLIVCTSGVDMPGADYQTLTKL 147

RESULT 4  
US-08-994-035C-5  
; Sequence 5, Application US/08994035C  
; Patent No. 6277625  
; GENERAL INFORMATION:  
; APPLICANT: Huang, Zhengyu  
; APPLICANT: Thomasow, Linda S  
; APPLICANT: Mavrodi, Dmitri V  
; APPLICANT: Raaijmakers, Jos M  
; APPLICANT: Weller, David M  
; APPLICANT: Cook, R James  
; TITLE OF INVENTION: Transgenic Strains for Biocontrol of  
; TITLE OF INVENTION: Plant Root Diseases  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Margaret A. Connor, Patent Advisor  
; STREET: 800 Buchanan St  
; CITY: Albany  
; STATE: CA



Query Match	15.7%;	Score 314;	DB 4;	Length 349;
Best Local Similarity	27.6%;	Pred. No. 7.4e-25;		
Matches 103;	Conservative 59;	Mismatched 169;	Indels 42;	Gaps 9;
30	CVEQSTYDPDFYFKTNS	---EHKTELKEKF----	ORMCDKSMIKRRYYWYLTEELKEN	80
5	CKPSLLFP--HYKIQQQMIDHLEQLHDDHPRMALAKRMIQNTQVNERIDVLVPIDELAVH	62		
81	PNVCEYMAPSLDARDQDWMVVEVPRIGCKEAAVKAIKEWGQPKSKITHLIVCTTSGVDMPGA	140		
63	TCGFTH-----RSIVYEREARRMSSIAARQAQIENAGLTTDDIRMVAVTSCGTGFWMPSL	114		
141	DYQLTKLGLRLPVYKRYMMYQGXFPAGGTVLRBLAKDLAENNKGARVLVVCSEVTATVFRG	200		
115	TAHLINDLGLRTSTVQLPQIAQLGCVAGAAAINRANDFASLSPDNHALIVLSFESSLCYQ-	173		
201	PSDTHLDSLVGOALPGDGNAAALIVGSDPVPEIEKEPIFEMVWTAQTAPDSEGAIDGHLRE	260		
174	PODTTKLHAFISAALFGDVAACWREADD----	KAPGFKIATGSGYFLPDSEHYKIVDKVD	229	
261	AGLTFLHLKDVPGIVSKNITKALVEAFELGHSIDVN-----SIFWIAHGGPAILDQVEQ	315		
230	SGFHTLTKAV-----MNSIKDVAPMVEELNFTETFNQHCQANDFFPHFGGRKILDELVL	284		
316	KLALKEPKKNATREVLSEYGNMSSACVLFIILDEMRRKSTQNGLTKTTGEGLEWGLVFGFGP	375		
285	OIDLPFGRVAOESDLSIENAGTISVVVFDVLKROFDSGPGANGAT-----GMLAAGFP	336		

RESULT 6  
US-08-434-907-12  
; Sequence 12, Application US/08494907  
; Patent No. 955298  
; GENERAL INFORMATION:  
; APPLICANT: Thomasow, Linda S  
; APPLICANT: Bargera, Mahalaxmi  
; APPLICANT: Weiler, David M  
; APPLICANT: Cook, R. James

;; TITLE OF INVENTION: Sequences for Production of  
;; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods  
;; NUMBER OF SEQUENCES: 20  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Margaret A. Connor, USDA-ARS  
;; STREET: 800 Buchanan Street  
;; CITY: Albany  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94710

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/494,907  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Connor, Margaret A  
;; REGISTRATION NUMBER: 30043  
;; REFERENCE/DOCKET NUMBER: 0009.95  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (510) 559-6067  
;; TELEFAX: (510) 559-5777  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 349 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; PCT-US96-10986-12

Query Match 15.5%; Score 310; DB 2; Length 349;  
Best Local Similarity 27.3%; Pred. No. 2e-24;  
Matches 102; Conservative 59; Mismatches 170; Indels 42; Gaps 9;

QY 30 CVEQSTYDPDFYFKITNS---EHKTELKEKF-----QRMCDKSMIKRYMYLTELKEN 80

Db 5 CKPSLLFP--HYKITQQQMHLEQLHDDHPRMALAKRMQNTQVNERLYLVIDELAVH 62

QY 81 PNVCYMAPSLDARQDMVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPGA 140

Db 63 TGFTTH-----RSIVYERARRMSSIAARQAIENAGLTDDIRVAVTSCGFMPSL 114

QY 141 DYQTLKLLGLRPVVKRYMYQGXFGAGTVLRKDLAENNKGARVLVVCSEVTAFTFG 200

Db 115 TAHLINDLGLRTSTVQLPIAQLGCVAGAAAINRANDFGSLSPDNHALIVSLFSSLCYQ- 173

QY 201 PSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMVWTAQTAPDSEGAIDGHLRE 260

Db 174 PQDTKLHAFISALFGDAVSACVWRADD---KAPGFKIAKTGSYFLPDSEHYIKYDVKD 229

QY 261 AGLTFLHLKDVPGIVSKNITKALVEAFEPGLIGSDYN-----SIFWIAHPGGPAILDQVEQ 315

Db 230 SGHFHTLDKAV-----MNSIKDVAPMMEELNFTFNQHCANDFFIHTGGRKILDELVL 284

QY 316 KLAKEPKKNATREVLSEYGNMSSACVLFIIDEMRKKTQNGLKTGEGLEWGLVFGFGP 375

Db 285 QLDLEPGVQAQRSDLSSEAGNIASVVVFDVLRQFDSGPANGAT-----GMLAAGFP 336

QY 376 GLTIETVTLRSVA 388

Db 337 GFTAEMAVGKWA 349

RESULT 7

PCT-US96-10986-12

;; Sequence 12, Application PC/TUS9610986

;; GENERAL INFORMATION:

;; TITLE OF INVENTION: Sequences for Production of

;; TITLE OF INVENTION: 2,4-Diacetylphloroglucinol and Methods

;; NUMBER OF SEQUENCES: 20  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Stephan A. Pendorf, DOMINIK & STEIN  
;; STREET: 600 N. West Shore Boulevard, Suite 1000  
;; CITY: Tampa  
;; STATE: FL  
;; COUNTRY: USA  
;; ZIP: 33609

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US96/10986  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Pendorf, Stephan A.  
;; REGISTRATION NUMBER: 32665  
;; REFERENCE/DOCKET NUMBER: A700.320  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (813) 289-2966  
;; TELEFAX: (813) 289-2967  
;; INFORMATION FOR SEQ ID NO: 12:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 349 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; PCT-US96-10986-12

Query Match 15.5%; Score 310; DB 5; Length 349;  
Best Local Similarity 27.3%; Pred. No. 2e-24;  
Matches 102; Conservative 59; Mismatches 170; Indels 42; Gaps 9;

QY 30 CVEQSTYDPDFYFKITNS---EHKTELKEKF-----QRMCDKSMIKRYMYLTELKEN 80

Db 5 CKPSLLFP--HYKITQQQMHLEQLHDDHPRMALAKRMQNTQVNERLYLVIDELAVH 62

QY 81 PNVCYMAPSLDARQDMVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPGA 140

Db 63 TGFTTH-----RSIVYERARRMSSIAARQAIENAGLTDDIRVAVTSCGFMPSL 114

QY 141 DYQTLKLLGLRPVVKRYMYQGXFGAGTVLRKDLAENNKGARVLVVCSEVTAFTFG 200

Db 115 TAHLINDLGLRTSTVQLPIAQLGCVAGAAAINRANDFGSLSPDNHALIVSLFSSLCYQ- 173

QY 201 PSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMVWTAQTAPDSEGAIDGHLRE 260

Db 174 PQDTKLHAFISALFGDAVSACVWRADD---KAPGFKIAKTGSYFLPDSEHYIKYDVKD 229

QY 261 AGLTFLHLKDVPGIVSKNITKALVEAFEPGLIGSDYN-----SIFWIAHPGGPAILDQVEQ 315

Db 230 SGHFHTLDKAV-----MNSIKDVAPMMEELNFTFNQHCANDFFIHTGGRKILDELVL 284

QY 316 KLAKEPKKNATREVLSEYGNMSSACVLFIIDEMRKKTQNGLKTGEGLEWGLVFGFGP 375

Db 285 QLDLEPGVQAQRSDLSSEAGNIASVVVFDVLRQFDSGPANGAT-----GMLAAGFP 336

QY 376 GLTIETVTLRSVA 388

Db 337 GFTAEMAVGKWA 349

RESULT 8

US-09-902-540-13778

;; Sequence 13778, Application US/09902540

;; Patent No. 6833447

;; GENERAL INFORMATION:

;; APPLICANT: Goldman, Barry S.

;; APPLICANT: Hinkle, Gregory J.

;; APPLICANT: Slater, Steven C.

**US-09-877-476-20**

Query Match            9.2%; Score 183; DB 4; Length 506;  
Best Local Similarity 23.4%; Pred. No. 1.6e-10;  
Matches         94; Conservative      63; Mismatches 159; Indels    86; Gaps    18

**Qy          27 PANC-VEQSYDPFYFKI-----TNSEHKTELKEKFMCDKS-----MKRR-----Y 69**  
**| | :| :**    |||||  
**| | :| :**    |||||

**Dd          88 PPCHRSVSKVMDFYQIRKADTSSRNGT-----CDNSSWLDFLRKIQERSGLGDE 138**  
**| | :| :**    |||||

**Qy          70 MYLTTEILKENPNVCYMARSLSARQMNVVEPRLCKEAAVKAIKWGPQSKIHLIV 129**  
**| | :| :**    |||||

**Dd          139 THQPEGILOVPPR--KTFFAARSETEOVIIGALENLFKNTV-----NPKD-Igilw 188**  
**| | :| :**    |||||

Qy 130 CTTSGVDMPGADYQLTKLLGLRPVVKRYMYMQGXKFGAGTTLVLAKOLAEENKKGARVLW 189  
:: : : : : : : : : : : : : : : : :  
Db 189 NSSNFNPPTSLSAMVVNTFKLSNVRSFNLGGMGCSAGVIAIDLAKOLLHVHKNTYALW 248  
Qy 190 CSEWTAWTFRGPSDTHLDLSVGQALFDGGAALIVGSDPVPPEISKPTFEMVWTQTATAPD 249

Qy	190	CSEVTA	TFRG	SDTH	LDLS	VGQA	LFGD	GAAL	IVGS	DPVPE	TEKPI	PMVW	TQA	TIAP	249
Db	249	STE--	-NITY	YAGD	NRSM	VMVSN	CLFR	VGGA	AILSN	KP-GD	RRSK	VELH	TVRT--	---	301

QY		250	SEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT-----	280
		302	UNCIIDYSERCTTCOCDFNCYCVCVSISVDITIVACBTWKNTATIGTIDPSIKLIFVV	361

QY 281 -----KALVAEPPLGTSIDNSIF--WIAHPGPAILDQVEQKLAKPKMNATREVLSE 333  
          :       :       :       :  
          :       :       :       :  
          :       :       :       :

Db 362 TFMGKPLPKDKIKHYVPDFKLAIDHFCIHAGGRAVIDVLEKNLALAPIDVEASRSTLHR 422

Qy 334 YGNMSSACVLFIIDEMRKKS-TQNLKTTGEGLEWGVLFQFG 374

Db  
422 FGN TSS SIW E L A Y I E A K R M K K G N V ----- W Q I A L G S G 457

RESULT 10  
US-09-877-476-16  
; Sequence 16, Application US/09877476

: PATENT NO.: 6713664  
 : GENERAL INFORMATION:  
 : APPLICANT: Jaworski, Jan G.  
 : APPLICANT: Blacklock, Brenda J.

```
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA  
; TITLE OF INVENTION: SYNTHASE POLYPEPTIDES  
; FILE REFERENCE: 07148-108001  
; FILE REFERENCE: 07148-108001  
; FILE REFERENCE: 07148-108001
```

CURRENT APPLICATION NUMBER: US 09/781,416  
 CURRENT FILING DATE: 2001-06-08  
 PRIOR APPLICATION NUMBER: US 60/210,326  
 PRIOR FILING DATE: 2000-06-08

```

; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```

```
:  
LENGTH: 506  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
:
```

OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAE1 (SEQ ID NO:2) and 3' 392 amino acids from B. napus elongase KCS (SEQ ID NO:4) having mutation at

US-09-877-476-16

OTHER INFORMATION: residue 307; designated AT114 G307D; hypotnetical

Query Match	9.1%	Score 181	DB 4	Length 506
-------------	------	-----------	------	------------

Best Local Similarity 22.9%; Pred. No. 2, 6e-10;  
Matches 95; Conservative 63; Mismatches 159; Indels 98; Gaps 18

QY	26	NNPANCVEOSTY	----	PDFYFKI	-----	INSHKTELKRFQMCDAS	-----	83
D6	75	NPVVLVDYSCLYPHPLKVSVMKDIFYQIRKADTSRRNGT	----	: : :	----	----	----	125

Qy	64	--MIKR-----YMLTTEILKENPVCYMPSLDARODMVVVEPRIGKEAAVKAIKE	111
		: : : : : :	
Db	126	LRKIQERSGLDGETHGPEGLQVPFR--KTFFAAREETEQIIGALENLFONTNV----	178

QY 117 WGPKSKITHLIVCTTSGVDMPCADYQLTLLGLRPVVKRYMMYQQXFGAGTVLR LAKD 176  
DB 179 --NPKD-IGILVNSNMFNPTSLSAMVNTFKLRNVRSFNLGGMGCSAGVIAIDLAKD 235  
QY 177 LAENNKARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETEKEPI 236  
DB 236 LLHVKNVTALVSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLSNKP-GDRRRSK 292  
QY 237 FEMVWTAQTIAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT- 280  
DB 293 YELVHTVRT-----HTGADDKSFRVCVQGGDENGKIGVLSKDI TDVAGRTVKKNIA TLGP 348  
QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGGPAILDQVEOKLAK 320  
DB 349 LILPLESEKLLFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDLEKNLALA 408  
QY 321 PERONATRELVSEYGNMSSACVLFILDENRKS--TONGLKTGTGEGLEWGLFGFG 374  
DB 409 PIDVEASRSTLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 11

US-09-877-476-22

; Sequence 22, Application US/09877476

; Patent No. 6713664

; GENERAL INFORMATION:

; APPLICANT: Jaworski, Jan G.

; APPLICANT: Blacklock, Brenda J.

; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,326

; PRIOR FILING DATE: 2000-06-08

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 22

; LENGTH: 506

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: 5' 114 amino acids from A. thaliana FAEL (SEQ ID

; OTHER INFORMATION: No:2) and 3' 392 amino acids from B. napus

; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having mutations at

; OTHER INFORMATION: positions 92 and 307; designated At14 K92R G307D;

; OTHER INFORMATION: hypothetical

US-09-877-476-22

Query Match 9.1%; Score 181; DB 4; Length 506;

Best Local Similarity 22.9%; Pred. No. 2.6e-10; Indels 98; Gaps 18;

Matches 95; Conservative 63; Mismatches 159;

QY 26 NPANCVEQSTY-----PDFYFKI-----TNSHKTLEKFKFORMCDKS----- 63

DB 75 NPVLVDYSCYLPPLHRLVSVSKVMDIFVQIRKADTSSRNGT-----CODSSWLD 125

QY 64 --MIKER-----YMYLTEEILKENPNVCEYMAPSLDARQDMVVVEVRLGKEAAVKAKE 116

DB 126 LRKIQRSGLGDBETHGPEGLQVPPR--KTFAAARETEQVIGALENLFKNTNV----- 178

QY 117 WGPKSKITHLIVCTTSGVDMPCADYQLTLLGLRPVVKRYMMYQQXFGAGTVLR LAKD 176

DB 179 --NPKD-IGILVNSNMFNPTSLSAMVNTFKLRNVRSFNLGGMGCSAGVIAIDLAKD 235

QY 177 LAENNKARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETEKEPI 236

DB 236 LLHVKNVTALVSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLSNKP-GDRRRSK 292

QY 237 FEMVWTAQTIAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT- 280

DB 293 YELVHTVRT-----HTGADDKSFRVCVQGGDENGKIGVLSKDI TDVAGRTVKKNIA TLGP 348

QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGGPAILDQVEOKLAK 320  
DB 349 LILPLESEKLLFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDLEKNLALA 408  
QY 321 PERONATRELVSEYGNMSSACVLFILDENRKS--TONGLKTGTGEGLEWGLFGFG 374  
DB 409 PIDVEASRSTLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 12

US-09-877-476-40

; Sequence 40, Application US/09877476

; Patent No. 6713664

; GENERAL INFORMATION:

; APPLICANT: Jaworski, Jan G.

; APPLICANT: Blacklock, Brenda J.

; TITLE OF INVENTION: FATTY ACID ELONGASE 3-KETOACYL COA

; FILE REFERENCE: 07148-108001

; CURRENT APPLICATION NUMBER: US/09/877,476

; CURRENT FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: US 60/210,326

; PRIOR FILING DATE: 2000-06-08

; NUMBER OF SEQ ID NOS: 56

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 40

; LENGTH: 506

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: 5' 173 amino acids from A. thaliana FAEL (SEQ ID

; OTHER INFORMATION: No:2) and 3' 333 amino acids from B. napus

; OTHER INFORMATION: elongase KCS (SEQ ID NO:4) having a mutation at

; OTHER INFORMATION: residue 307; designated At173 G307D; hypothetical

US-09-877-476-40

Query Match 9.0%; Score 180; DB 4; Length 506;

Best Local Similarity 23.2%; Pred. No. 3.3e-10; Indels 80; Gaps 17;

Matches 94; Conservative 62; Mismatches 170;

QY 26 NPANCVEQSTY-PDFYFKITNSE-----HKTLEKFKFORMCDKS-----MIKRR-- 68

DB 75 NPVLVDYSCYLPPLHRLVSVSKVMDIFVQIRKADTSSRNVACDDPSSLDLFLRKIQERSG 134

QY 69 ---YMYLTEEILKENPNVCEYMAPSLDARQDMVVVEVRLGKEAAVKAKEWQPKSKIT 125

DB 135 LGDETYSPEGLIHVPPR--KTFAASRETEKVIIGALENLFKNTNV-----NPKD-IG 184

QY 126 HLIVCTTSGVDMPCADYQLTLLGLRPVVKRYMMYQQXFGAGTVLR LAKDLAENNKAR 185

DB 185 ILVNSSMFNPTSLSAMVNTFKLRNVRSFNLGGMGCSAGVIAIDLAKOLLHVHKNTY 244

QY 186 VLVVCEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDPVPETEKEPIFEMVWTAQT 245

DB 245 ALWVSTE--NITYNIYAGDNRMVSNCLFRVGGAAILLSNKP-GDRRSKVELVHTVRT 301

QY 246 IAPDSEGAIDGHLR-----EAG-LTFHLLKDVPGI-----VSKNIT----- 280

DB 302 ---HTGADDKSFRVCVQGGDENGKIGVLSKDI TDVAGRTVKKNIA TLGPLILPLSEKL 357

QY 281 -----KALVEAFEPGLGSDYNSIF--WIAHPGGPAILDQVEOKLAKPEKNMAYRE 329

DB 358 LFFVTMGKLLFKDKIKHYVVPDFKLAIDHFCIHAGGRAVIDLEKNLALAPIDVEASRS 417

QY 330 VLSEYGNMSSACVLFILDENRKS--TONGLKTGTGEGLEWGLFGFG 374

DB 418 TLHRFGNTSSSIWYELAYIEAKGRMKGNKV-----WQIALGSG 457

## RESULT 13

US-08-657-749D-6

; Sequence 6, Application US/08657749D



```

Query Match      8.9%; Score 178; DB 4; Length 505;
Best Local Similarity 22.6%; Pred. No. 5.4e-10;
Matches 95; Conservative 63; Mismatches 170; Indels 92; Gaps 18;

QY 18 TILAIGT-ANPANCVEQSY-----PDYFKITNSEHKTKEKPFQMCCKS 63
Db 66 SVLYIATRPKPYLVVEYSCYLPPTCHRSSISKVMDIFFQVRKADPSR-----NGTCDDSD 119
QY 64 -----MIKRR-----YMLTEIILKENPNVCEYNAPSILDAQDMVVVEVRLGKEAAV 111
Db 120 SWLDFLRKIQERGLGDETHGPEGLQVPPR--KTFARARETEQVILGALENLFKNTNV 177
QY 112 KAIKRGQPKSKITHLIVCTTSCGDMFGADYQLTKLLGLRPYVKRYMYTQGXFGAGTVL 171
Db 178 -----NPKD-IGILVNSSMENPTPSLSAMVNTFKLSNVRSENLGGMGCSAGVIAI 229
QY 172 RLAKOLAENKKGARVLVWCSEVTAFTFRGSDTHLDSLVGQALFGDGAALIVGSDPVEE 231
Db 230 DLAKDLLHVHKNYALVWSTE--NITYIYAGNRSMMVSNCLFRVGGAAILLNKP-RD 286
QY 232 IEKPIFEMVWTAQTIAPDSEGAIDGHLR-----EAGLT-PHLLKDVPGI-----VSKNI 279
Db 287 RRSKYELVHTVRT-----HTGADKSFRCVQOGDDENGQTVSLSKDITDVAGRTVKNI 342
QY 280 T-----KALVEAFEPGLIGISDYNISF--WIAHPGGPAILDOVEQ 315
Db 343 ATGLPLILPLSEKLLFFVTFMGKLFKDEIKHYVDFPKLAIDHFCIHAGGKAVIDVLEK 402
QY 316 KLALKEKNATREVLSEYGNMSSACVLFILDEMCKS--TQNGLKTGTGEGWGLFGFG 374
Db 403 NLGLAPIDVEASRSTLHRTGNTSSSIWYELAYIEPKGRMKGNKV-----WQIALGSG 456

```

Search completed: April 22, 2005, 21:45:54  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:34:56 ; Search time 40 Seconds  
(without alignments)  
935.708 Million cell updates/sec

Title: US-10-031-918A-1  
Perfect score: 1997  
Sequence: 1 MVSVEIRKQRAEGPATIL.....LFGFGGLTIETVLRSAI 389

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1994	99.8	389	2 S35164	naringenin-chalcon
2	1969	98.6	389	2 S44370	naringenin-chalcon
3	1943	97.3	383	2 S35165	naringenin-chalcon
4	1941	97.2	389	2 S35166	naringenin-chalcon
5	1938	97.0	389	2 S26414	naringenin-chalcon
6	1932	96.7	389	2 S20932	naringenin-chalcon
7	1926	96.4	389	2 S35167	naringenin-chalcon
8	1926	96.4	389	2 S33210	naringenin-chalcon
9	1915	95.9	389	2 S20933	naringenin-chalcon
10	1911	95.7	389	2 S44367	naringenin-chalcon
11	1880	94.1	389	2 S35163	naringenin-chalcon
12	1875	93.9	370	2 S44368	naringenin-chalcon
13	1867	93.5	389	1 SYFJCP	naringenin-chalcon
14	1839	92.1	389	2 JQ2250	naringenin-chalcon
15	1818	91.0	389	2 S49202	naringenin-chalcon
16	1814	90.8	389	2 S49203	naringenin-chalcon
17	1811.5	90.7	388	1 SY5YC1	naringenin-chalcon
18	1809.5	90.6	388	2 JQ2249	naringenin-chalcon
19	1805.5	90.4	388	2 S60472	naringenin-chalcon
20	1800.5	90.2	388	1 SY5YC3	naringenin-chalcon
21	1800.5	90.2	388	2 S37098	naringenin-chalcon
22	1795.5	89.9	388	1 SY5YC1	naringenin-chalcon
23	1794.5	89.9	388	2 JQ2259	naringenin-chalcon
24	1767	88.5	389	1 SYFJCP	naringenin-chalcon
25	1755	87.9	389	2 JCS136	naringenin-chalcon
26	1735	86.9	390	1 SY5KCD	naringenin-chalcon
27	1734	86.8	389	1 SYFJCP	naringenin-chalcon
28	1731	86.7	389	1 SYFJCP	naringenin-chalcon
29	1720	86.1	389	2 T07799	naringenin-chalcon

30 1705 85.4 396 2 S20515 naringenin-chalcon  
31 1703.5 85.3 388 2 JCS516 naringenin-chalcon  
32 1697 85.0 410 2 S12224 naringenin-chalcon  
33 1695.5 84.9 394 1 SYJCS naringenin-chalcon  
34 1680 84.1 398 2 S56699 naringenin-chalcon  
35 1674.5 83.9 395 1 SYISC3 naringenin-chalcon  
36 1667.5 83.5 395 1 SYMUCN naringenin-chalcon  
37 1660.5 83.1 395 1 SYMUCN naringenin-chalcon  
38 1658 83.0 381 2 S12223 naringenin-chalcon  
39 1652 82.7 398 2 S42523 naringenin-chalcon  
40 1648 82.5 403 2 S55464 chalcone synthase  
41 1637 82.0 391 2 T10713 naringenin-chalcon  
42 1634 81.8 398 2 S58190 naringenin-chalcon  
43 1632 81.7 400 1 SYZMCC naringenin-chalcon  
44 1614 80.8 398 2 S16275 naringenin-chalcon  
45 1575 78.9 331 2 JQ1071 naringenin-chalcon

ALIGNMENTS

RESULT 1  
S35164  
naringenin-chalcon synthase (EC 2.3.1.74) 2 - alfalfa  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S35164  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and exp.  
A:Reference number: S35163; MUID:93283629; PMID:8507827  
A:Accession: S35164  
A:Molecule type: mRNA  
A:Residues: 1-389 <JUN>  
A:Cross-references: UNIPROT:P30074; EMBL:L02902; NID:g166363; PIDN:AAA02824.1; PID:g166363  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 99.8%; Score 1994; DB 2; Length 389;  
Best Local Similarity 99.7%; Pred. No. 4.7e-143;  
Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVEIRKQRAEGPATILAIQTANPANCVEQSTYPPDFYFKITNSEHKTELKEKFORMC 60  
DB 1 MVSVEIRKQRAEGPATILAIQTANPANCVEQSTYPPDFYFKITNSEHKTELKEKFORMC 60  
QY 61 DKSMIKRRYMYLTETILKENPNVCYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRRYMYLTETILKENPNVCYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPYVKRYMYQCGXFGTGLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMFGADYQTLKGLRPYVKRYMYQCGXFGTGLAKDLAEN 180  
QY 181 NKGARLVVCSVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPEIEKPIFENV 240  
DB 181 NKGARLVVCSVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPGIVSKNITKALVEAPEPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTTHLLKDVPGIVSKNITKALVEAPEPLGSDYNSIFW 300  
QY 301 IAHGCGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKSTQNGLKT 360  
DB 301 IAHGCGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKSTQNGLKT 360  
QY 361 TGEGLGWGLFGFGPLGTIETVLRSAI 389  
DB 361 TGEGLGWGLFGFGPLGTIETVLRSAI 389

RESULT 2

S44370  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44370  
R:McKhann, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa)  
A:Reference number: S44367; MUID:94250839; PMID:8193301  
A:Accession: S44370  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-389 <MCK>  
A:Cross-references: UNIPROT:P30075; EMBL:U01021; NID:g393000; PIDN:AAB41559.1; PID:g393000  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A

Query Match 98.6%; Score 1969; DB 2; Length 389;  
Best Local Similarity 98.5%; Pred. No. 3.6e-141;  
Matches 383; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
QY 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTATFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
DB 181 NKGARVLVVCSEVTATFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240  
QY 241 WTAQTAPSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
DB 241 WTAQTAPSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALPKPERKNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKT 360  
DB 301 IAHPGGPAILDQVEQKALPKPERKNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKT 360  
QY 361 TGEGLRWGLVFGFGPGGLTIETVVLRSVAI 389  
DB 361 TGEGLRWGLVFGFGPGGLTIETVVLRSVAI 389

RESULT 3  
S35165  
naringenin-chalcone synthase (EC 2.3.1.74) 4 - alfalfa (fragment)  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 15-Mar-2004  
C:Accession: S35165; S36949  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and expression of the cDNA for the chalcone synthase gene  
A:Reference number: S35163; MUID:93283629; PMID:8507827  
A:Accession: S35165  
A:Molecule type: mRNA  
A:Residues: 1-331 <JUN>  
A:Cross-references: EMBL:L02903  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
submitted to the EMBL Data Library, September 1992  
A:Reference number: S36949  
A:Accession: S36949  
A:Molecule type: mRNA  
A:Residues: 53-383 <JU2>  
A:Cross-references: EMBL:L02903; NID:g166365; PIDN:AAA02825.1; PID:g166366

C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 97.3%; Score 1943; DB 2; Length 383;  
Best Local Similarity 98.4%; Pred. No. 3.3e-139;  
Matches 377; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 7 IRKQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMCDSKMIK 66  
DB 1 IRKQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMCDSKMIK 60  
QY 67 RRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQPKSKITH 126  
DB 61 RRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQPKSKITH 120  
QY 127 LIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAENKKGARV 186  
DB 121 LIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAENKKGARV 180  
QY 187 LVVCSEVTATFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMVMTAQT 246  
DB 181 LVVCSEVTATFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMVMTAQT 240  
QY 247 APDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFWIAHPGG 306  
DB 241 APDSEGAIDGHLREAGLTFLHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFWIAHPGG 300  
QY 307 PAILDQVEQKALPKPERKNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKTGEGLE 366  
DB 301 PAILDQVEQKALPKPERKNATREVLSEYGNMSSACVLFILDMRKKSTQNGLKTGEGLE 360  
QY 367 WGLVFGFGPGGLTIETVVLRSVAI 389  
DB 361 WGLVFGFGPGGLTIETVVLRSVAI 383

## RESULT 4

S35166  
naringenin-chalcone synthase (EC 2.3.1.74) 8 - alfalfa  
N:Alternate names: chalcone synthase  
C:Species: Medicago sativa (alfalfa)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C:Accession: S35166  
R:Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A:Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and expression of the cDNA for the chalcone synthase gene  
A:Reference number: S35163; MUID:93283629; PMID:8507827  
A:Accession: S35166  
A:Molecule type: mRNA  
A:Residues: 1-389 <JUN>  
A:Cross-references: UNIPROT:P30076; EMBL:L02904; NID:g166367; PIDN:AAA02826.1; PID:g166368  
A:Note: the authors translated the codon GTT for residue 344 as Phe  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 97.2%; Score 1941; DB 2; Length 389;  
Best Local Similarity 97.4%; Pred. No. 4.8e-139;  
Matches 379; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
DB 1 MVSVEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKKEKFORMC 60  
QY 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLITEILKENPNVCEYMAPSLDARDQMVVVEVPRLGKEAAVKAKEWGQ 120  
QY 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
DB 121 KSKITHLIVCTTSGVDMFGADYQTLKLLGRPVYKRYMYQQGCFAGGTVLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVTATFRGSPDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFEMV 240



Db 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNITKALVEAFEPPLGIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNITKALVEAFEPPLGIGSDYNSIFW 300  
Qy 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Db 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
Db 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389

## RESULT 5

S26414  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C;Species: Medicago sativa (alfalfa)  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
R;Accession: S26414  
R;Buffard, D.; Breda, C.; El Turk, J.; Sallaud, C.; Kondorosi, A.; Esnault, R.  
submitted to the EMBL Data Library, August 1992  
A;Description: Molecular cloning of two chalcone synthase cDNA from alfalfa.  
A;Reference number: S26414  
A;Accession: S26414  
A;Molecule type: mRNA  
A;Residues: 1-389 <BUF>  
A;Cross-references: UNIPROT:P51078; EMBL:X69106; NID:g19590; PIDN:CAA48226.1; PID:g19591  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 97.0%; Score 1938; DB 2; Length 389;  
Best Local Similarity 97.2%; Pred. No. 8e-139; 9; Indels 0; Gaps 0;  
Matches 378; Conservative 2; Mismatches 9;  
Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Db 1 MVSVEIRNAQRAEGPATILAI GTANPTNCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Qy 61 DKSMIKRRYMYLTERILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTERILKENPSVCEIMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Db 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIKPIFEMV 240  
Db 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNITKALVEAFEPPLGIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNITKALVEAFEPPLGIGSDYNSIFW 300  
Qy 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Db 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
Db 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389

## RESULT 6

S20932  
naringenin-chalcone synthase (EC 2.3.1.74) 2 - garden pea  
N;Alternate names: chalcone synthase  
C;Species: Pisum sativum (garden pea)  
C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
R;Accession: S20932  
R;Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.

Plant Mol. Biol. 18, 1009-1012, 1992  
A;Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A;Reference number: S20931; MUID:92256801; PMID:1581561  
A;Accession: S20932  
A;Molecule type: mRNA  
A;Residues: 1-389 <ICH>  
A;Cross-references: UNIPROT:Q01287; EMBL:X63334; NID:g20825; PIDN:CAA44934.1; PID:g20826  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 96.7%; Score 1932; DB 2; Length 389;  
Best Local Similarity 96.4%; Pred. No. 2.3e-138;  
Matches 375; Conservative 6; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Db 1 MVTVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Qy 61 DKSMIKRRYMYLTERILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTERILKENPSLCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Qy 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Db 121 KSKITHLIVCTTSGVDMPGADYQLTLLGLRPYVKRYMYQQCXFAGGTVLRKDLAEN 180  
Qy 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIKPIFEMV 240  
Db 181 NKNARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPPIEIKPIFEMV 240  
Qy 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNITKALVEAFEPPLGIGSDYNSIFW 300  
Db 241 WTAQTIAPDSEGAIDGHLRAGLTFFHLKDVPGIVSKNIDKALVEAFKPLGIGSDYNSIFW 300  
Qy 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Db 301 IAHPPGPAILDQVEQKALPEKXNATREVLSEYGNMSSACVLFILDEMRRKKSQTQGLTKT 360  
Qy 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
Db 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389

## RESULT 7

S35167  
naringenin-chalcone synthase (EC 2.3.1.74) 9 - alfalfa  
N;Alternate names: chalcone synthase  
C;Species: Medicago sativa (alfalfa)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
R;Accession: S35167  
R;Junghane, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A;Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and exp  
A;Reference number: S35163; MUID:93283629; PMID:8507827  
A;Accession: S35167  
A;Molecule type: mRNA  
A;Residues: 1-389 <JUN>  
A;Cross-references: UNIPROT:P30077; EMBL:L02905; NID:g166369; PIDN:AAA02827.1; PID:g166  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 96.4%; Score 1926; DB 2; Length 389;  
Best Local Similarity 95.9%; Pred. No. 6.5e-138;  
Matches 373; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MVSVEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Db 1 MVSVEIRQAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKQRMVC 60  
Qy 61 DKSMIKRRYMYLTERILKENPNVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120  
Db 61 DKSMIKRRYMYLTERILKENPSVCEYMAPSLDARDQMDVVEVPRLGKEAAVKAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
DB 121 KSKITHLICTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLDWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 8  
S33610  
naringenin-chalcone synthase (EC 2.3.1.74) 1 - garden pea  
N:Alternate names: chalcone synthase  
C:Species: Pisum sativum (garden pea)  
C>Date: 08-Dec-1993 #sequence\_revision 23-Feb-1996 #text\_change 09-Jul-2004  
C:Accession: S33610, S20931  
R:An, C.; Ichinose, Y.; Yamada, T.; Tanaka, Y.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 21, 789-803, 1993  
A:Title: Organization of the genes encoding chalcone synthase in Pisum sativum.  
A:Reference number: S33610; MUID:93222478; PMID:8467077  
A:Accession: S33610  
A:Molecule type: DNA  
A:Residues: 1-389 <ANC>  
A:Cross-references: UNIPROT:Q01286; EMBL:D10661; NID:g391780; PID:g3917  
R:Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 18, 1009-1012, 1992  
A:Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A:Reference number: S20931; MUID:92256801; PMID:1581561  
A:Accession: S20931  
A:Molecule type: mRNA  
A:Residues: 1-389 <ICH>  
A:Cross-references: EMBL:X63333; NID:g20823; PID:CAA44933.1; PID:g20824  
C:Genetics:  
A:Introns: 60/1  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis  
Query Match 96.4%; Score 1926; DB 2; Length 389;  
Best Local Similarity 96.1%; Pred. No. 6 5e-138;  
Matches 374; Conservative 5; Mismatches 10; Indels 0; Gaps 0;  
QY 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVSIRKQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRYYLITEILKENPNVCEYMAPSLDARDQDMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRYYLITEILKENPNVCEYMAPSLDARDQDMVVVEVPRLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLDWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 10  
S44367  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44367  
R:McKham, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa).  
A:Reference number: S44367; MUID:94250839; PMID:8193301  
A:Accession: S44367  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA

QY 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 9  
S20933  
naringenin-chalcone synthase (EC 2.3.1.74) 3 - garden pea  
N:Alternate names: chalcone synthase  
C:Species: Pisum sativum (garden pea)  
C>Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S20933  
R:Ichinose, Y.; Kawamata, S.; Yamada, T.; An, C.; Kajiwara, T.; Shiraishi, T.; Oku, H.  
Plant Mol. Biol. 18, 1009-1012, 1992  
A:Title: Molecular cloning of chalcone synthase cDNAs from Pisum sativum.  
A:Reference number: S20931; MUID:92256801; PMID:1581561  
A:Accession: S20933  
A:Molecule type: mRNA  
A:Residues: 1-389 <ICH>  
A:Cross-references: UNIPROT:Q01288; EMBL:X63335; NID:g20827; PID:CAA44935.1; PID:g20828  
C:Superfamily: Type III polyketide synthase  
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis  
Query Match 95.9%; Score 1915; DB 2; Length 389;  
Best Local Similarity 95.4%; Pred. No. 4 4e-137;  
Matches 371; Conservative 6; Mismatches 12; Indels 0; Gaps 0;  
QY 1 MVSVSIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
DB 1 MVSVSIRKAQRAEGPATILAI GTATPANCVEQSTYDPFYFKITNSEHKTTELKEKFORMC 60  
QY 61 DKSMIKRYYLITEILKENPNVCEYMAPSLDARDQDMVVVEVPRLGKEAAVKAKEWGQP 120  
DB 61 DKSMIKRYYLITEILKENPNVCEYMAPSLDARDQDMVVVEVPRLGKEAAVKAKEWGQP 120  
QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
DB 121 KSKITHLIFCTTSGVDMPGADYQTLKGLRPPVKKRYMMYQOQXFGAGTGLRLAKDLAEN 180  
QY 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
DB 181 NKGARVLVVCSEVAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVPPEIEKPIFENV 240  
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVPGIVSKNITKALVEAFPLGSDYNSIFW 300  
QY 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPGGPAILDQVEQKALKEPKKMATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
QY 361 TGEGLWGVLFPGFPGGLTIETVVLRSVAI 389  
DB 361 TGEGLDWGVLFPGFPGGLTIETVVLHSVAI 389  
RESULT 10  
S44367  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S44367  
R:McKham, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A:Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa).  
A:Reference number: S44367; MUID:94250839; PMID:8193301  
A:Accession: S44367  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA

A;Residues: 1-389 <MCK>  
A;Cross-references: UNIPROT:P51077; EMBL:U01018; NID:G392994; PIDN:AAB41561.1; PID:G392994  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 95.7%; Score 1911; DB 2; Length 389;  
Best Local Similarity 94.9%; Pred. No. 8.8e-137;  
Matches 363; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVSSEIRKQRAEGPATILAIATGTPANPCVEOSTYDPDFYFKITNSEHKTELKEKQRM 60  
DB 1 MVSSEIRKQRAEGPATILAIATGTPANPCVEOSTYDPDFYFKITNSEHKTELKEKQRM 60

QY 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQ 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAEN 180

QY 181 NKGARLVVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
DB 181 NKGARLVVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFW 300

QY 301 IAHPPGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPPGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389

RESULT 11  
S35163  
naringenin-chalcone synthase (EC 2.3.1.74) 1 - alfalfa  
N;Alternate names: chalcone synthase  
C;Species: Medicago sativa (alfalfa)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 09-Jul-2004  
C;Accession: S35163  
R;Jungmans, H.; Dalkin, K.; Dixon, R.A.  
Plant Mol. Biol. 22, 239-253, 1993  
A;Title: Stress responses in alfalfa (Medicago sativa L.). 15. Characterization and expression of the cDNA for the chalcone synthase gene  
A;Reference number: S35163; MUID:93283629; PMID:8507827  
A;Molecule type: mRNA  
A;Residues: 1-389 <JUN>  
A;Cross-references: UNIPROT:P30073; EMBL:L02901; NID:G166361; PIDN:AAA02823.1; PID:G166361  
A;Note: the authors translated the codon ATT for residue 284 as Val  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 94.1%; Score 1880; DB 2; Length 389;  
Best Local Similarity 93.3%; Pred. No. 1.9e-134;  
Matches 363; Conservative 9; Mismatches 17; Indels 0; Gaps 0;

QY 1 MVSSEIRKQRAEGPATILAIATGTPANPCVEOSTYDPDFYFKITNSEHKTELKEKQRM 60  
DB 1 MVSSEIRKQRAEGPATILAIATGTPANPCVEOSTYDPDFYFKITNSEHKTELKEKQRM 60

QY 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQ 120  
DB 61 DKSMIKRRYMYLVEELKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQ 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAEN 180  
DB 121 KSKITHLIVCTTSGVDMPGADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAEN 180

QY 181 NKGARLVVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240  
DB 181 NKGARLVVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFW 300  
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFW 300

QY 301 IAHPPGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360  
DB 301 IAHPPGPAILDQVEOKLALPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389  
DB 361 TGEGLWGVLFPGFGPLTIETVVLRSVAI 389

RESULT 12  
S44368  
naringenin-chalcone synthase (EC 2.3.1.74) - alfalfa  
C;Species: Medicago sativa (alfalfa)  
C;Date: 13-Jan-1995 #sequence\_revision 10-Nov-1995 #text\_change 15-Mar-2004  
C;Accession: S44368  
R;McKhann, H.I.; Hirsch, A.M.  
Plant Mol. Biol. 24, 767-777, 1994  
A;Title: Isolation of chalcone synthase and chalcone isomerase cDNAs from alfalfa (Medicago sativa L.)  
A;Reference number: S44367; MUID:94250839; PMID:8193301  
A;Accession: S44368  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-370 <MCK>  
A;Cross-references: EMBL:U01019; NID:G392996; PIDN:AAB41560.1; PID:G514830  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1993  
C;Superfamily: Type III polyketide synthase  
C;Keywords: acyltransferase; coenzyme A

Query Match 93.9%; Score 1875; DB 2; Length 370;  
Best Local Similarity 98.1%; Pred. No. 4.3e-134;  
Matches 363; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 20 LAIGTANPANCVEOSTYDPDFYFKITNSEHKTELKEKQRMCDKSMIKRRYMYLVEELKE 79  
DB 1 LAIGTANPANCVEOSTYDPDFYFKITNSEHKTELKEKQRMCDKSMIKRRYMYLVEELKE 79

QY 80 NPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPG 139  
DB 61 NPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQPKSKITHLIVCTTSGVDMPG 120

QY 140 ADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAENNKARLVVVCSEVTAVTFR 199  
DB 121 ADYQTLKLLGLRPYVKRYMYQQXFPAGGTVLRKOLAENNKARLVVVCSEVTAVTFR 180

QY 200 GPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMVWTAQTIAPDSEGAIDGHLR 259  
DB 181 GPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFEMVWTAQTIAPDSEGAIDGHLR 240

QY 260 EAGLTFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFWIAHPGGPAILDQVEOKLAL 319  
DB 241 EAGLTFHLLKDVPGIVSKNITKALVEAPEPLGIGSDYNSIFWIAHPGGPAILDQVEOKLAL 300

QY 320 KPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKTGEGLEWGVLFPGFGPLTI 379  
DB 301 KPEKKNATREVLSEYGNMSSACVLFILDEMRRKSTONGLKTGEGLEWGVLFPGFGPLTI 360

QY 380 ETWVLRSVAI 389  
DB 361 ETWVLRSVAI 370

RESULT 13  
SYFJCP  
naringenin-chalcone synthase (EC 2.3.1.74) I - kudzu vine

QY 181 NKGARVLVVCSEVTAVTFRGPDTHLDSLVGQALFGDGAALI VGS DVPPEIEKDFEMV 240

[illegible]

Search completed: April 22, 2005, 21:45:05  
Job time : 44 secs

1103

608.01(A)

2163.07—  
(b)

2173.05(S)

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2005, 21:25:54 ; Search time 165 Seconds  
(without alignments)

911.817 Million cell updates/sec

Title: US-10-031-918A-1

Perfect score: 1997

Sequence: 1 MVSVEIRKAQRAEGPATIL.....LFGFGGLTETVLRVAI 389

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 segs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: Geneseq1980s:\*

2: Geneseq1990s:\*

3: Geneseq2000s:\*

4: Geneseq2001s:\*

5: Geneseq2002s:\*

6: Geneseq2003as:\*

7: Geneseq2003bs:\*

8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1995	99.9	389	4	AAB68589 Chalcone
2	1994	99.8	389	5	ABG31832 Polyketid
3	1991	99.7	389	5	ABG31809 Polyketid
4	1991	99.7	389	5	ABG31804 Polyketid
5	1991	99.7	389	5	ABG31802 Polyketid
6	1991	99.7	389	5	ABG31817 Polyketid
7	1991	99.7	389	5	ABG31806 Polyketid
8	1990	99.6	389	5	ABG31821 Polyketid
9	1990	99.6	389	5	ABG31805 Polyketid
10	1990	99.6	389	5	ABG31813 Polyketid
11	1990	99.6	389	5	ABG31803 Polyketid
12	1990	99.6	389	5	ABG31812 Polyketid
13	1988	99.5	389	5	ABG31819 Polyketid
14	1988	99.5	389	5	ABG31808 Polyketid
15	1987	99.5	389	5	ABG31820 Polyketid
16	1987	99.5	389	5	ABG31822 Polyketid
17	1987	99.5	389	5	ABG31815 Polyketid
18	1986	99.4	389	5	ABG31807 Polyketid
19	1986	99.4	389	5	ABG31818 Polyketid
20	1986	99.4	389	5	ABG31801 Polyketid
21	1986	99.4	389	5	ABG31810 Polyketid
22	1986	99.4	389	5	ABG31811 Polyketid
23	1985	99.4	389	5	ABG31816 Polyketid
24	1969	98.6	389	8	ADM32822 Amino aci
25	1933	96.8	389	6	ABU94275 Trifolium

#### ALIGNMENTS

##### RESULT 1

AAB68589

ID AAB68589 standard; protein; 389 AA.

XX AC AAB68589;

XX DT 24-APR-2001 (first entry)

XX DE Chalcone synthase.

XX KW Chalcone synthase; enzyme; protein coordinate data; alfalfa;

XX KW polyketide synthase; plant phenylpropanoid biosynthesis; antibiotic;

XX KW anticancer agent; antifungal agent.

XX OS Medicago sativa.

XX PN WO200107579-A2.

XX PD 01-FEB-2001.

XX PF 27-JUL-2000; 2000WO-US020674.

XX PR 27-JUL-1999; 99US-0145898P.

XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Ferrer J, Jez J, Austin M, Bowman M;

XX WPI; 2001-159712/16.

XX New isolated polyketide synthases, useful for identifying substrates and inhibitors and for producing potential pharmaceuticals.

XX Disclosure; Page 12; 213pp; English.

XX The present sequence is chalcone synthase (CHS) from Medicago sativa (alfalfa), which is a polyketide synthase. The chalcone synthase has at least 14 active site alpha-carbon atoms with structural coordinates defined in the specification. CHS plays an essential role in the biosynthesis of plant phenylpropanoids. Crystalline forms of CHS are used to determine the three-dimensional co-ordinates of polyketide synthases (PKS), especially of the active site, and from this information potential substrates and inhibitors of PKS can be modelled and tested, also predictions may be made about activity and substrate specificity of putative PKS. The co-ordinates may also be used to design mutant PKS that can produce different and/or new polyketide compounds, potentially useful as antibiotics, anticancer and antifungal agents

XX

SQ Sequence 389 AA;

Query Match 99.9%; Score 1995; DB 4; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 6e-203;  
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGOP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTA VTFRGPSTDLHDSLVGQALFGDGAALIVGSDPVPETKEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTA VTFRGPSTDLHDSLVGQALFGDGAALIVGSDPVPETKEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYNISIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYNISIFW 300

QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSQTQGLKT 360

QY 361 TGEGLWGVLFPGFGPLTITVLRSAI 389  
 DB 361 TGEGLWGVLFPGFGPLTITVLRSAI 389

RESULT 2  
 ABG31832  
 ID ABG31832 standard; protein; 389 AA.  
 XX ABG31832;  
 XX 05-NOV-2002 (first entry)  
 DT Polyketide synthase.  
 DE Polyketide synthase.  
 XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure.  
 XX Mammalia.  
 XX WO200257418-A2.  
 XX 25-JUL-2002.  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX 15-DEC-2000; 2000US-0255811P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noel.JP, Austin MB, Bowman ME;  
 XX WPI; 2002-590730/63.  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 PT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX Claim 2; Page 13; 243pp; English.  
 XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS

CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone or in the formation of both  
 CC resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents wild-type PKS amino acid sequence used to make the PKS mutants  
 CC of the invention  
 XX

SQ Sequence 389 AA;

Query Match 99.8%; Score 1994; DB 5; Length 389;  
 Best Local Similarity 99.7%; Pred. No. 7.7e-203;  
 Matches 388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPDFYFKITNSEHKTELKEKFORMC 60

QY 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGOP 120  
 DB 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARDQMDVVVEVPRLGKEAAVKAKEWGOP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGRLPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180

QY 181 NKGARVLVVCSEVTA VTFRGPSTDLHDSLVGQALFGDGAALIVGSDPVPETKEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTA VTFRGPSTDLHDSLVGQALFGDGAALIVGSDPVPETKEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYNISIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFELGISDYNISIFW 300

QY 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEOKLALKEPKMNATREVLSEYGNMSSACVLFILDEMRKKSQTQGLKT 360

QY 361 TGEGLWGVLFPGFGPLTITVLRSAI 389  
 DB 361 TGEGLWGVLFPGFGPLTITVLRSAI 389

RESULT 3  
 ABG31809  
 ID ABG31809 standard; protein; 389 AA.  
 XX ABG31809;  
 XX 05-NOV-2002 (first entry)  
 DT Polyketide synthase mutant #9.  
 DE Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutein.  
 XX Mammalia.  
 OS Synthetic.  
 XX Key Location/Qualifiers  
 FH





SQ Sequence 389 AA;  
 Query Match 99.7%; Score 1991; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTELKEKFORMC 60  
 QY 61 DKSMTKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRGLKEAAVKAKEWGQP 120  
 DB 61 DKSMTKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRGLKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGLRPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGLRPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKEPKMNA TREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKMNA TREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 QY 361 TGEGLWGVLFPGFGPLGTIETV LRSVAI 389  
 DB 361 TGEGLWGVLFPGFGPLGTIETV LRSVAI 389  
 RESULT 5  
 ABG31802  
 ID ABG31802 standard; protein; 389 AA.  
 AC ABG31802;  
 XX 05-NOV-2002 (first entry)  
 XX Polyketide synthase mutant #2.  
 XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 XX chalcone; crystal structure; mutant; mutin.  
 OS Mammalia.  
 OS Synthetic.  
 PH Key Location/Qualifiers  
 FT Misc-difference 98 /note= "Wild type Val substituted by Leu"  
 FT  
 XX WO200257418-A2.  
 XX 25-JUL-2002.  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX 15-DEC-2000; 2000US-0255811P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX Noël JP, Austin MB, Bowman ME;  
 XX WPI; 2002-590730/63.  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 XX of wild-type (wt) PKS with crystal structure of second PKS and  
 XX substituting amino acids of wt PKS with amino acids at homologous  
 XX positions in second PKS.

Claim 4; Page: 243pp; English.  
 The invention relates to a method of producing mutant polyketide synthase (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS with a crystal structure of a second PKS; (b) substituting one or more amino acids of the wild-type PKS with the amino acid residues at homologous positions in the second PKS; and (c) producing the mutant PKS. The method is useful for producing a mutant polyketide synthase, and for altering the activity of PKS, where the altered activity results in the formation of the product of the second PKS instead of the product of the wild-type PKS. Preferably, the altered activity results in the formation of resveratrol instead of chalcone or in the formation of both resveratrol and chalcone. The crystalline form of PKS is useful for determining the position of specific alpha-carbon atoms and R-groups associated with it, comprising the active site, in three-dimensional space. It is also suitable for X-ray or neutron diffraction analysis to determine three-dimensional structure of mutant PKS and to design additional mutants. The crystallisation can serve as further purification. Because the synthase may crystallise in more than one crystal form, the structural coordinates of alpha-carbon atoms of an active site determined from a synthase or its portions are useful to solve structure of other crystal forms of synthases. The present sequence represents a mutant polyketide synthase of the invention. Note: The present sequence is not shown in the specification but was derived from the wild-type PKS sequence given on page 13 (see ABG31832)  
 SQ Sequence 389 AA;  
 Query Match 99.7%; Score 1991; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAI GTANPANCVEQSTYDPFYFKITNSEHKTELKEKFORMC 60  
 QY 61 DKSMTKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRGLKEAAVKAKEWGQP 120  
 DB 61 DKSMTKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVEVPRGLKEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQLT KLGLRPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQLT KLGLRPYVKRYMYQQGXFGAGTVLRLAKDLAEN 180  
 QY 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSTHLDLSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLT FHLKDVPGIVSKNITKALVEAFPLGISDYSNIFW 300  
 QY 301 IAHPGGPAILDQVEQKALKEPKMNA TREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKEPKMNA TREVLSEYGNMSSACVLFILDEMRRKKSQTQGLKT 360  
 QY 361 TGEGLWGVLFPGFGPLGTIETV LRSVAI 389  
 DB 361 TGEGLWGVLFPGFGPLGTIETV LRSVAI 389  
 RESULT 6  
 ABG31817  
 ID ABG31817 standard; protein; 389 AA.  
 AC ABG31817;  
 XX 05-NOV-2002 (first entry)  
 XX Polyketide synthase mutant #17.  
 XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;



CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.7%; Score 1991; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 1.6e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSSEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60

QY 61 DKSMIKRMYLTYEELTKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAIAKEWGQP 120  
 DB 61 DKSMIKRMYLTYEELTKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAIAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQQGXFGAGTIVLRKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQQGXFGAGTIVLRKDLAEN 180

QY 181 NKGARVLVVCSEVTATVTRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATVTRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLGTIETVLRSAI 389  
 DB 361 TGEGLWGVLFPGFGPLGTIETVLRSAI 389

RESULT 8  
 ABG31821  
 ID ABG31821 standard; protein; 389 AA.  
 AC ABG31821;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Polyketide synthase mutant #21.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; reeseratrol;  
 KW chalcone; crystal structure; mutant; muten.  
 XX  
 OS Mammalia.  
 OS Synthetic.  
 XX  
 FH \* Key Location/Qualifiers  
 FT Misc-difference 270  
 FT /note= "Wild type Asp substituted by Ala"  
 XX  
 PN WO200257418-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 14-DEC-2001; 2001WO-US048523.  
 XX  
 PR 15-DEC-2000; 2000US-0255811P.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Noel JP, Austin MB, Bowman ME;  
 XX

DR WPI; 2002-590730/63.  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 FT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX

PS Claim 4; Page; 243pp; English.

XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of reeseratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)

SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;

Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSSEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60  
 DB 1 MVSSEIRKAQRAEGPATILAIQTANPANCVEQSTYDPDFYFKITNSEHKTTELKEKFORMC 60

QY 61 DKSMIKRMYLTYEELTKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAIAKEWGQP 120  
 DB 61 DKSMIKRMYLTYEELTKENPNVCEYMAPSLDARQDMVVVEVPRLGKAAVKAIAKEWGQP 120

QY 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQQGXFGAGTIVLRKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLTKLGLRPYVKRYMYQQGXFGAGTIVLRKDLAEN 180

QY 181 NKGARVLVVCSEVTATVTRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTATVTRGSDTHLDSLVGOALFGDGAALIVGSDPVPEIEKPIFEMV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFLHLLKDVFGIVSKNITKALVEAFELGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALKPEKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGPLGTIETVLRSAI 389  
 DB 361 TGEGLWGVLFPGFGPLGTIETVLRSAI 389

RESULT 9  
 ABG31805  
 ID ABG31805 standard; protein; 389 AA.

XX ABG31805;

XX

DT 05-NOV-2002 (first entry)  
 DE Polyketide synthase mutant #5.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutein.  
 XX  
 OS Mammalia.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 131 /note= "Wild type Thr substituted by Ser"  
 FT  
 XX  
 PN WO200257418-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 14-DEC-2001; 2001WO-US049523.  
 XX  
 PR 15-DEC-2000; 2000US-0255811P.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Noel JP, Austin MB, Bowman ME;  
 XX  
 DR WPI; 2002-590730/63.  
 XX  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 PT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX  
 PS Claim 4; Page; 243pp; English.  
 XX  
 XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31813)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVEIRKQRAEGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTELKEKQRM 60  
 DB 1 MVSVEIRKQRAEGPATILAIATNPANCVEQSTYDPDFYFKITNSEHKTELKEKQRM 60  
 QY 61 DKSMIKRRYMYLTELKENPNVCCEYMAPSLDARQDMVVVEVPRLGEAAVKAKEWGQP 120  
 DB 61 DKSMIKRRYMYLTELKENPNVCCEYMAPSLDARQDMVVVEVPRLGEAAVKAKEWGQP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQGCFFAGGTVLRKOLAEN 180

DB 121 KSKITHLIVCTTSGVDMPGADYQLTKLLGLRPYVKRYMMYQGCFFAGGTVLRKOLAEN 180  
 QY 181 NKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 DB 181 NKGARVLVVCSEVTAATFRGSDTHLDSLVGOALFGDGAALIVGSDPVPPIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPKPMNATREVLSEYGNMSSACVLFILDEMRKSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALPKPMNATREVLSEYGNMSSACVLFILDEMRKSTONGLKT 360  
 QY 361 TGEGLWGVLFPGPGLTITVLRSAI 389  
 DB 361 TGEGLWGVLFPGPGLTITVLRSAI 389  
 RESULT 10  
 ABG31813  
 ID ABG31813 standard; protein; 389 AA.  
 XX  
 AC ABG31813;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Polyketide synthase mutant #13.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutein.  
 XX  
 OS Mammalia.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 160 /note= "Wild type Tyr substituted by Phe"  
 FT  
 XX  
 PN WO200257418-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 14-DEC-2001; 2001WO-US048523.  
 XX  
 PR 15-DEC-2000; 2000US-0255811P.  
 XX  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX  
 PI Noel JP, Austin MB, Bowman ME;  
 XX  
 DR WPI; 2002-590730/63.  
 XX  
 XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 PT of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.  
 XX  
 PS Claim 4; Page; 243pp; English.  
 XX  
 XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31813)  
 XX  
 SQ Sequence 389 AA;

CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKs and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKs sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAQRAEGPATILGATNPANCVQSTYDPDFYFKITNSEHKTLEKQFQMC 60  
 DB 1 MVSVEIRKAQRAEGPATILGATNPANCVQSTYDPDFYFKITNSEHKTLEKQFQMC 60  
 QY 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWQOP 120  
 DB 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWQOP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTGLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTGLRLAKDLAEN 180  
 QY 181 NGARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDVPPEIEKPIFEMV 240  
 DB 181 NGARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDVPPEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGVLFQFGPGLTIETVWLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVWLRSAI 389

RESULT 11  
 ABG31803  
 ID ABG31803 standard; protein; 389 AA.  
 XX  
 AC ABG31803;  
 XX  
 DT 05-NOV-2002 (first entry)  
 XX  
 DE Polyketide synthase mutant #3.  
 XX  
 KW Polyketide synthase; enzyme; protein coordinate data; PKs; resveratrol;  
 XX chalcone; crystal structure; mutant; mutein.  
 OS Mammalia.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 99 /note= "Wild type Val substituted by Ala"  
 FT  
 XX W0200257418-A2.  
 XX  
 XX 25-JUL-2002.  
 XX  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX  
 XX 15-DEC-2000; 2000US-025581P.  
 PR

XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PA Noel JP, Austin MB, Bowman ME;  
 PI WPI; 2002-590730/63.  
 XX  
 DR Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 XX of wild-type (wt) PKs with crystal structure of second PKs and  
 PT substituting amino acids of wt PKs with amino acids at homologous  
 PT positions in second PKs.  
 XX  
 PS Claim 4; Page; 243pp; English.  
 XX  
 CC The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKs  
 CC with a crystal structure of a second PKs; (b) substituting one or more  
 CC amino acids of the wild-type PKs with the amino acid residues at  
 CC homologous positions in the second PKs; and (c) producing the mutant PKs.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKs, where the altered activity results in the  
 CC formation of the product of the second PKs instead of the product of the  
 CC wild-type PKs. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKs is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKs and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKs sequence given on page 13 (see ABG31832)  
 XX  
 SQ Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 2e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MVSVEIRKAQRAEGPATILGATNPANCVQSTYDPDFYFKITNSEHKTLEKQFQMC 60  
 DB 1 MVSVEIRKAQRAEGPATILGATNPANCVQSTYDPDFYFKITNSEHKTLEKQFQMC 60  
 QY 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWQOP 120  
 DB 61 DKSMIKRRYMYLTHEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWQOP 120  
 QY 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTGLRLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPGADYQTLKGLRYPYVRYMYQOQXFGAGTGLRLAKDLAEN 180  
 QY 181 NGARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDVPPEIEKPIFEMV 240  
 DB 181 NGARVLVVCSEVTAFTFRGSDTHLDSLVGOALFGDGAALIVGSDVPPEIEKPIFEMV 240  
 QY 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFFHLKDVFGIVSKNITKALVEAFPLGSDYNSIFW 300  
 QY 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 DB 301 IAHPGGPAILDQVEQKALPKPEKNATREVLSEYGNMSSACVLFILDEMRRKSTQNGLKT 360  
 QY 361 TGEGLWGVLFQFGPGLTIETVWLRSAI 389  
 DB 361 TGEGLWGVLFQFGPGLTIETVWLRSAI 389

RESULT 12



```

ABG31812
ID ABG31812 standard; protein; 389 AA.
XX AC
XX ABG31812;
XX DT
XX 05-NOV-2002 (first entry)
XX DE
XX Polyketide synthase mutant #12.
XX KW
XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;
XX chalcone; crystal structure; mutant; mutein.
XX KW
XX Mammalia.
XX OS
XX Synthetic.
XX FH
XX Key
XX FT Misc-difference 159
XX /note= "Wild type Met substituted by Val"
XX PN
XX WO200257418-A2.
XX XX
XX 25-JUL-2002.
XX PF
XX 14-DEC-2001; 2001WO-US048523.
XX PR
XX 15-DEC-2000; 2000US-0255811P.
XX PA
XX (SALK ) SALK INST BIOLOGICAL STUDIES.
XX PI
XX Noel JP, Austin MB, Bowman ME;
XX WPI; 2002-590730/63.
XX DR
XX
XX PT
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure
XX of wild-type (wt) PKS with crystal structure of second PKS and
XX substituting amino acids of wt PKS with amino acids at homologous
XX positions in second PKS.
XX PS
XX Claim 4; Page; 243pp; English.
XX CC
XX The invention relates to a method of producing mutant polyketide synthase
XX (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS
XX with a crystal structure of a second PKS; (b) substituting one or more
XX amino acids of the wild-type PKS with the amino acid residues at
XX homologous positions in the second PKS; and (c) producing the mutant PKS.
XX The method is useful for producing a mutant polyketide synthase, and for
XX altering the activity of PKS, where the altered activity results in the
XX formation of the product of the second PKS instead of the product of the
XX wild-type PKS. Preferably, the altered activity results in the formation
XX of resveratrol instead of chalcone. The crystalline form of PKS is useful
XX for determining the position of specific alpha-carbon atoms and R-groups
XX associated with it, comprising the active site, in three-dimensional
XX space. It is also suitable for x-ray or neutron diffraction analysis to
XX determine three-dimensional structure of mutant PKS and to design
XX additional mutants. The crystallisation can serve as further
XX purification. Because the synthase may crystallise in more than one
XX crystal form, the structural coordinates of alpha-carbon atoms of an
XX active site determined from a synthase or its portions are useful to
XX solve structure of other crystal forms of synthases. The present sequence
XX represents a mutant polyketide synthase of the invention. Note: The
XX present sequence is not shown in the specification but was derived from
XX the wild-type PKS sequence given on page 13 (see ABG31832)
XX SQ
XX Sequence 389 AA;

Query Match 99.6%; Score 1990; DB 5; Length 389;
Best Local Similarity 99.5%; Pred. No. 2e-202;
Matches 387; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MVSVSIRKQRAEGPATIIAIGTANPANCVEQSTYPDFYFKITNSEHKTELKEKQRM 60
DB 1 MVSVSIRKQRAEGPATIIAIGTANPANCVEQSTYPDFYFKITNSEHKTELKEKQRM 60

```

```

QY 61 DKSMIKRRYMYLTTEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120
DB 61 DKSMIKRRYMYLTTEEILKENPNVCEYMAPSLDARQDMVVVEVPRLGKEAAVKAKEWGQP 120
QY 121 KSKITHLIVCTTSGVDMPGADYQLTKLGLRPVVKRYMYVOQGXFAGGTVLRKAKDLAEN 180
DB 121 KSKITHLIVCTTSGVDMPGADYQLTKLGLRPVVKRYMYVOQGXFAGGTVLRKAKDLAEN 180
QY 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240
DB 181 NKGARVLVVCSEVTAVTFRGSDTHLDSLVGQALFGDGAALIVGSDPVPEIEKPIFEMV 240
QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFBPLGIDSINSIFW 300
DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVFGIVSKNITKALVEAFBPLGIDSINSIFW 300
QY 301 IAHPGGPAILDQVEOKLALPKPMNATREVLSEYGNMSSACVLFIIDEMRKSTQNGLKT 360
DB 301 IAHPGGPAILDQVEOKLALPKPMNATREVLSEYGNMSSACVLFIIDEMRKSTQNGLKT 360
QY 361 TGEGLWGVLFPGFGLTITVTLRSVAI 389
DB 361 TGEGLWGVLFPGFGLTITVTLRSVAI 389

RESULT 13
ABG31819
ID ABG31819 standard; protein; 389 AA.
XX AC
XX ABG31819;
XX DT
XX 05-NOV-2002 (first entry)
XX DE
XX Polyketide synthase mutant #19.
XX KW
XX Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;
XX chalcone; crystal structure; mutant; mutein.
XX OS
XX Mammalia.
XX OS
XX Synthetic.
XX FH
XX Key
XX Location/Qualifiers
XX FT Misc-difference 268
XX /note= "Wild type Leu substituted by Lys"
XX PN
XX WO200257418-A2.
XX PD
XX 25-JUL-2002.
XX PF
XX 14-DEC-2001; 2001WO-US048523.
XX PR
XX 15-DEC-2000; 2000US-0255811P.
XX PA
XX (SALK ) SALK INST BIOLOGICAL STUDIES.
XX PI
XX Noel JP, Austin MB, Bowman ME;
XX WPI; 2002-590730/63.
XX DR
XX
XX PT
XX Producing mutant polyketide synthase (PKS) by comparing crystal structure
XX of wild-type (wt) PKS with crystal structure of second PKS and
XX substituting amino acids of wt PKS with amino acids at homologous
XX positions in second PKS.
XX PS
XX Claim 4; Page; 243pp; English.
XX CC
XX The invention relates to a method of producing mutant polyketide synthase
XX (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS
XX with a crystal structure of a second PKS; (b) substituting one or more
XX amino acids of the wild-type PKS with the amino acid residues at
XX homologous positions in the second PKS; and (c) producing the mutant PKS.
XX The method is useful for producing a mutant polyketide synthase, and for
XX altering the activity of PKS, where the altered activity results in the
XX formation of the product of the second PKS instead of the product of the
XX wild-type PKS. Preferably, the altered activity results in the formation
XX of resveratrol instead of chalcone. The crystalline form of PKS is useful
XX for determining the position of specific alpha-carbon atoms and R-groups
XX associated with it, comprising the active site, in three-dimensional
XX space. It is also suitable for x-ray or neutron diffraction analysis to
XX determine three-dimensional structure of mutant PKS and to design
XX additional mutants. The crystallisation can serve as further
XX purification. Because the synthase may crystallise in more than one
XX crystal form, the structural coordinates of alpha-carbon atoms of an
XX active site determined from a synthase or its portions are useful to
XX solve structure of other crystal forms of synthases. The present sequence
XX represents a mutant polyketide synthase of the invention. Note: The
XX present sequence is not shown in the specification but was derived from
XX the wild-type PKS sequence given on page 13 (see ABG31832)
XX SQ
XX Sequence 389 AA;

```

CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)

XX SQ Sequence 389 AA;

Query Match 99.5%; Score 1988; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 3.3e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVSIRKAQRAEGPATILAICTANPANCVEQSTYDFYFKITNSEHKTLEKFORMC 60  
 DB 1 MVSVSIRKAQRAEGPATILAICTANPANCVEQSTYDFYFKITNSEHKTLEKFORMC 60

QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGP 120  
 DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGP 120

QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKRYMYQQCXFAGGTVLAKDLAEN 180  
 DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKRYMYQQCXFAGGTVLAKDLAEN 180

QY 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFENV 240  
 DB 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300  
 DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEOKLALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360  
 DB 301 IAHPGGPAILDQVEOKLALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389  
 DB 361 TGEGLWGVLFPGFGLTIETVLRSAI 389

RESULT 14

ABG31808  
 ID ABG31808 standard; protein; 389 AA.

AC ABG31808;

XX 05-NOV-2002 (first entry)

DE Polyketide synthase mutant #8.

KW Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;  
 KW chalcone; crystal structure; mutant; mutein.

XX Mammalia.  
 OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 135 /note= "Wild type Val substituted by Pro"

XX WO200257418-A2.

PD 25-JUL-2002.  
 XX 14-DEC-2001; 2001WO-US048523.  
 XX 15-DEC-2000; 2000US-0255811P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.

XX Noel JP, Austin MB, Bowman MB;  
 PI WPI; 2002-590730/63.

XX Producing mutant polyketide synthase (PKS) by comparing crystal structure  
 of wild-type (wt) PKS with crystal structure of second PKS and  
 PT substituting amino acids of wt PKS with amino acids at homologous  
 PT positions in second PKS.

XX Claim 4; Page; 243pp; English.

XX The invention relates to a method of producing mutant polyketide synthase  
 CC (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS  
 CC with a crystal structure of a second PKS; (b) substituting one or more  
 CC amino acids of the wild-type PKS with the amino acid residues at  
 CC homologous positions in the second PKS; and (c) producing the mutant PKS.  
 CC The method is useful for producing a mutant polyketide synthase, and for  
 CC altering the activity of PKS, where the altered activity results in the  
 CC formation of the product of the second PKS instead of the product of the  
 CC wild-type PKS. Preferably, the altered activity results in the formation  
 CC of resveratrol instead of chalcone. The crystalline form of PKS is useful  
 CC for determining the position of specific alpha-carbon atoms and R-groups  
 CC associated with it, comprising the active site, in three-dimensional  
 CC space. It is also suitable for X-ray or neutron diffraction analysis to  
 CC determine three-dimensional structure of mutant PKS and to design  
 CC additional mutants. The crystallisation can serve as further  
 CC purification. Because the synthase may crystallise in more than one  
 CC crystal form, the structural coordinates of alpha-carbon atoms of an  
 CC active site determined from a synthase or its portions are useful to  
 CC solve structure of other crystal forms of synthases. The present sequence  
 CC represents a mutant polyketide synthase of the invention. Note: The  
 CC present sequence is not shown in the specification but was derived from  
 CC the wild-type PKS sequence given on page 13 (see ABG31832)

XX SQ Sequence 389 AA;

Query Match 99.5%; Score 1988; DB 5; Length 389;

Best Local Similarity 99.5%; Pred. No. 3.3e-202;

Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MVSVSIRKAQRAEGPATILAICTANPANCVEQSTYDFYFKITNSEHKTLEKFORMC 60

DB 1 MVSVSIRKAQRAEGPATILAICTANPANCVEQSTYDFYFKITNSEHKTLEKFORMC 60

QY 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGP 120

DB 61 DKSMIKRMYLITEILKENPNVCEYMAPSLDARQDMVVEVPRIGKEAAVKAKEWGP 120

QY 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKRYMYQQCXFAGGTVLAKDLAEN 180

DB 121 KSKITHLIVCTTSGVDMPCADYQTLKGLRPVVKRYMYQQCXFAGGTVLAKDLAEN 180

QY 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFENV 240

DB 181 NKGARVLVVCSEVTAVTFRGPSDTHLDSLVGQALFGDGAALIVGSDPVEIEKPIFENV 240

QY 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300

DB 241 WTAQTIAPDSEGAIDGHLREAGLTFHLLKDVPGIVSKNITKALVEAFEPGLGSDYNSIFW 300

QY 301 IAHPGGPAILDQVEOKLALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

DB 301 IAHPGGPAILDQVEOKLALKEPKKNATREVLSEYGNMSSACVLFILDEMRKKSSTONGLKT 360

QY 361 TGEGLWGVLFPGFGLTIETVLRSAI 389



```

Db      361 TGEGLWGLFGFGPGLTIETVLRVAI 389
|||||
RESULT 15
ABG31820
ID      ABG31820 standard; protein; 389 AA.
XX      AC      ABG31820;
XX      AC      05-NOV-2002 (first entry)
XX      DE      Polyketide synthase mutant #20.
XX      KW      Polyketide synthase; enzyme; protein coordinate data; PKS; resveratrol;
XX      KW      chalcone; crystal structure; mutant; muten.
XX      OS      Mammalia.
XX      OS      Synthetic.
XX      FH      Key      Location/Qualifiers
XX      FT      Misc-difference 269
XX      FT      /note= "Wild type Lys substituted by Gly"
XX      PN      WO200257418-A2.
XX      XX
XX      PD      25-JUL-2002.
XX      XX
XX      PF      14-DEC-2001; 2001WO-US048523.
XX      XX
XX      PR      15-DEC-2000; 2000US-0255811P.
XX      XX
XX      PA      (SALK ) SALK INST BIOLOGICAL STUDIES.
XX      XX
XX      PI      Noel JP, Austin MB, Bowman ME;
XX      XX
XX      DR      WPI; 2002-590730/63.
XX      XX
XX      PT      Producing mutant polyketide synthase (PKS) by comparing crystal structure
XX      PT      of wild-type (wt) PKS with crystal structure of second PKS and
XX      PT      substituting amino acids of wt PKS with amino acids at homologous
XX      PT      positions in second PKS.
XX      XX
XX      PS      Claim 4; Page; 243pp; English.
XX      CC
XX      CC      The invention relates to a method of producing mutant polyketide synthase
XX      CC      (PKS) comprising: (a) comparing a crystal structure of a wild-type PKS
XX      CC      with a crystal structure of a second PKS; (b) substituting one or more
XX      CC      amino acids of the wild-type PKS with the amino acid residues at
XX      CC      homologous positions in the second PKS; and (c) producing the mutant PKS.
XX      CC      The method is useful for producing a mutant polyketide synthase, and for
XX      CC      altering the activity of PKS, where the altered activity results in the
XX      CC      formation of the product of the second PKS instead of the product of the
XX      CC      wild-type PKS. Preferably, the altered activity results in the formation
XX      CC      of resveratrol instead of chalcone. The crystalline form of PKS is useful
XX      CC      for determining the position of specific alpha-carbon atoms and R-groups
XX      CC      associated with it, comprising the active site, in three-dimensional
XX      CC      space. It is also suitable for X-ray or neutron diffraction analysis to
XX      CC      determine three-dimensional structure of mutant PKS and to design
XX      CC      additional mutants. The crystallisation can serve as further
XX      CC      purification. Because the synthase may crystallise in more than one
XX      CC      crystal form, the structural coordinates of alpha-carbon atoms of an
XX      CC      active site determined from a synthase or its portions are useful to
XX      CC      solve structure of other crystal forms of synthases. The present sequence
XX      CC      represents a mutant polyketide synthase of the invention. Note: The
XX      CC      present sequence is not shown in the specification but was derived from
XX      CC      the wild-type PKS sequence given on page 13 (see ABG31832)
XX      SQ      Sequence 389 AA;

```

Query Match 99.5%; Score 1987; DB 5; Length 389;  
 Best Local Similarity 99.5%; Pred. No. 4.3e-202;  
 Matches 387; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Search completed: April 22, 2005, 21:41:19  
 Job time : 167 secs

*This Page Blank (usbrn)*